Proposed IMMA Revisions

Revised DRAFT, 11 July 2012

S. Woodruff, E. Freeman, S. Lubker, S. Smith, S. Worley et al.

Introduction

Format changes—to yield IMMA version 1—are marked below in blue within Supplement C of the current IMMA (version 0) documentation. The abbreviated version of the documentation (http://icoads.noaa.gov/e-doc/imma/R2.5-imma_short.pdf) includes Supp. D, providing details on individual field configurations; whereas the complete documentation (http://icoads.noaa.gov/e-doc/imma/R2.5-imma.pdf) also provides additional background in its main text and in Supps. A-B, including more detailed comparisons between IMMA and international exchange (IMMT/IMMPC) formats.

Detailed background/discussion notes for the *Core* and attachments (attms) appear following each table below. When significant changes or additions to attm content were needed (i.e. for *Immt*, *Mod-qc*, and *Meta-vos*) a revised attm was created with a new table number and new attachment ID (*ATTI*), but the old configuration was retained as a deprecated attm. The *Ivad*, *Error*, and *Uid* attms are new, and the *Pocn*, *Auto*, *Track*, and *Hist* are proposed (with many details of proposed attms to be finalized; including field abbreviations, which need to be checked for uniqueness to avoid possible overlaps with already defined field abbreviations).

The following items discuss aspects of the planned changes in greater detail:

1. Abbreviated structural element names in italics: A new IMMA documentation feature—to enhance communication—is that the *Core* and attms (e.g. *Icoads*) are all given abbreviated names in *italics*, but with only the first letter in uppercase, to distinguish them from the fully capitalized IMMA field abbreviations.

2. Switch to attm-internal field numbering: E.g. within the *Icoads* attm, the fields are now numbered 1-51 rather than 49-99. Otherwise documentation maintenance was becoming challenging, and in conjunction with the IVAD project this is viewed as a more flexible approach. However, the revised Fortran program to read/write IMMA1 (rdimma1[1]) still utilizes a linear field numbering approach for assigning array storage across the *Core* and all (including deprecated) attms.

3. Operational and deprecated attms: Rather than change (i.e. add/subtract, or modify, fields) attms, a new attm version is created, tentatively with tables for the deprecated attms retained only in the last-version format documentation (and noting that by itself the item 2 field numbering switch is not considered such a change). With minor modifications however, the rdimma1 program is able to read (and write) both the operational and deprecated attms (note: extending even to storing both a new and deprecated attm within a given IMMA record), and the format remains fully backward compatible. Related also to item 2, the new *Ivad* and *Error* attms include both input component number (i.e. 0=*Core* or *ATTI*) (*ICN*) and field number (*FN*).

[1] http://icoads.noaa.gov/software/rdimma1. With limited modifications, the program can also write IMMA1. In terms of processing new attms, at present the program will process only the *Uid* attm; coding for the *Ivad* and *Error* attms is not yet implemented, also processing of Subsidiary records is not yet implemented. However, additional software to meet these requirements will be developed in conjunction with prototyping.

4. Additional software maintenance considerations: To make translation software adapted from rdimma (e.g. existing adaptations of rdimma0 used to translate data from other formats into IMMA) more robust over the longer term, usage of field abbreviations (e.g. FTRUE(SST)) rather than hard-coded storage array locations (e.g. FTRUE(35)) appears prudent, but could, for instance if programmers copy modifications into the new rdimma shell, possibly have unwanted/unexpected behavior with respect to referring to deprecated rather than currently operational attm fields (e.g. *W2* is defined in both Tables C2 and C5; note: in rdimma1 however, the fields name constant definitions associated with deprecated attms are by default commented out i.e. inoperative).

5. Switch to multi-record "linked" approach: Rather than modifying the *Icoads* attm as was originally proposed to include *UID* and associated release-tracking information etc., those fields are placed in a short new *Uid* attm (see Table C98), which appears both in the Main and (any optional) Subsidiary records, linking them all together (see further discussion following Table C8), e.g.:

Main IMMA record type: *Core* + *Icoads* + *Immt* + *Mod-qc* + *Meta-vos* + *Uid* + *Suppl*
Subsidiary IMMA record type: *Uid* + *Ivad* + *Ivad* + *Ivad* … + *Ivad*
Subsidiary IMMA record type: *Uid* + *Error* + *Error* + *Error* … + *Error*
Subsidiary IMMA record type: *Uid* + *Meta-vos*

The way this is being implemented is still anticipated to be fully backward compatible. The *Core* has not changed, but rdimma1 checks the first character of each record to determine if it is the Main (starts with "1" or "2") or a Subsidiary (starts with "9") record.

6. Anticipated software constraints on Main and Subsidiary records and their attm composition: Processing by rdimma1 of Subsidiary records, and of the *Ivad* and *Error* attms, is not yet implemented (as noted above). However, at least the following constraints should be feasible: (i) Subsidiary records must each begin with a *Uid* attm, followed by zero or more attms of any type except for the *Uid* type. (ii) The maximum number of *Ivad* and *Error* attms within a Subsidiary record tentatively will be set at 100 each, and probably no checks will be made for repeating attms (e.g. two *Ivad* attms referring to SST, in which case possibly the second attm would overwrite information from the first). (iii) Any other attms must be non-repeating, and any that overlap with Main would overwrite the information from Main. In addition, while by definition of *UID* two (or more) Main records should never appear with the same *UID*, no check is envisioned as feasible for *UID* uniqueness.

7. Status of new-field content: In the revised attms (*Immt*, *Mod-qc*, and *Meta-vos*) newly defined fields generally will not be populated (i.e. in the prototype Release 2.5 dataset), with the exception of field *MDS* in the *Meta-vos* attm (see also related discussion in Annex C). In the new *Uid* attm, *UID* will be set as discussed below Table C98, and the Release number fields will be set as *RN1*=2, *RN2*=5, *RN3*=1, i.e. this will be considered R2.5.1, to distinguish it from the original R2.5.

These Annexes provide additional background information:

- Annex A: Implementation/Unresolved Issues
- Annex B: Planned ICOADS Development of a Unique Report ID (*UID*); and Intra-record Release No. (*RN*) Tracking
- Annex C: Reprocessing Notes for Recovering Missing Field Configurations, Etc.
- Annex D: Scenarios for Adding Adjustments or QC within IVAD
- Annex E: Discussion of *Ivad* attm Configuration Details (w/ UK NOC)
- Annex F: QC Flag Discussion
- Annex G: Edited Cloud Report (ECR) Information

Supplement C. Record Types

The IMMA *Core* (Table C0) forms the common front-end for all record types. By itself, the Core, which is divided into location and regular sections, forms a useful abbreviated record type incorporating many of the most commonly used data elements in standardized form (drawn from the fields to be agreed internationally, listed in Supp. D). Concatenating one or more "attachments" (attm) after the *Core* creates additional record types. So far, in addition to the *Core*, the following attms have been defined:

Table C0: Core (*Core*)	(108 characters)
Table C1: ICOADS (*Icoads*) attm	(65 characters)
Table C5: IMMT-5/FM 13 (*Immt*) attm	(94 characters)
Table C6: Model quality control (*Mod-qc*) attm	(68 characters)
Table C7: Ship metadata (*Meta-vos*) attm	(58 characters)
Table C8: ICOADS Value-added Database (*Ivad*) attm	(implementation TBD)
Table C9: Error (*Error*) attm	(implementation TBD)
Table C98: Unique ID (*Uid*) attm	(15 characters)
Table C99: Supplemental data (*Suppl*) attm	(length may vary)

including these deprecated attms (note: not documented here; see IMMA0 documentation):

Table C2: IMMT-2/FM 13 attm	(76 characters)
Table C3: Model quality control attm	(66 characters)
Table C4: Ship metadata attm	(57 characters)

Additionally, the following attms are proposed (CP):

Table CP1: Physical oceanographic (*Pocn*) attm	(45 characters)
Table CP2: Automated instrumentation (*Auto*) attm	(41 characters)
Table CP3: Platform tracking (*Track*) attm	(proposed)
Table CP4: Historical (*Hist*) attm	(proposed)

and the following attms are envisioned as further possibilities, but without any suggested content below:

Buoy metadata (*Meta-buoy*) attm	(proposed, no table)
Reanalyses quality control (*Rean-qc*) attm	(proposed, no table)
Daily observational (*Daily*) attm	(possible, no table)

The following are examples of the record types that can be constructed from the Core plus these attachments (where Table numbers are used to indicate the corresponding attm) (NOTE: ~~strikethrough~~ below indicates updates are still needed, or marks material that has been moved elsewhere):

- Core:
  C0 (108 characters)
- ICOADS-standard structure (used for Release 2.5, see Supp. E):
  ~~C0 + C1 + C2 + C3 + C4 + C6 (372 characters, before C6)~~
- NCDC-variant structure (used alternatively for Release 2.5, see Supp. E):
  ~~C0 + C1 + C2 + C3 + C6 (315 characters, before C6)~~

• historical record:

~~C0 + C5 + C6 (proposed)~~

Inclusion of the attm count (*ATTC*) field in the Core, and of the attm ID (*ATTI*) and attm data length (*ATTL*) fields at the beginning of each attm, enables computer parsing of the records. Thus additional variations on these basic record types are implemented by inclusion or omission of attms, and new attms can be defined in the future as needed for new data or metadata requirements.

Each table following contains these columns:

1: Field number. Field numbering is attm-internal beginning with field number 1 and ending with the last field indicated in the table.
2: Length (Len.) in characters (i.e. bytes).
3-4: Abbreviation (Abbr.) for each element (or field), and a brief description.
5-6: For fields with a numeric range, the minimum (Min.) and maximum (Max.) are indicated. In other cases the range and configuration are listed as: "a" for alphabetic (A-Z), "b" for alphanumeric (strictly 0-Z), "c" for alphanumeric plus other characters, or "u" for undecided form (only for fields that are currently unused).
7: Units of data and related WMO Codes. Information in parentheses usually relates the proposed field to a field from Supp. B, Table B1 (if applicable): WMO Code symbolic letters are listed, or "•" followed by a field number from Table B1 in the absence of symbolic letters. This information is prefixed by "Δ" to highlight field configurations that are extended in range or modified in form from presently defined WMO representations.

Table C0. IMMA Core.

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
			Location section (45 characters):			
1	4	*YR*	year UTC	1600	2024	(AAAA)
2	2	*MO*	month UTC[1]	1	12	(MM)
3	2	*DY*	day UTC[1]	1	31	(YY)
4	4	*HR*	hour UTC[1]	0	23.99	0.01 hour (Δ GG)
5	5	*LAT*	latitude	–90.00	90.00	0.01°N (Δ $L_aL_aL_a$)
6	6	*LON*	longitude[1]	–179.99	359.99	0.01°E (Δ $L_oL_oL_oL_o$)
				0.00	359.99	(ICOADS convention)
				–179.99	180.00	(NCDC-variant convention)
7	2	*IM*	IMMA version	0	99	(Δ •65)
8	1	*ATTC*	attm count	0	9	
9	1	*TI*	time indicator	0	3	
10	1	*LI*	latitude/long. indic.	0	6	
11	1	*DS*	ship course	0	9	(D_s)
12	1	*VS*	ship speed	0	9	(Δ v_s)
13	2	*NID*	national source indic.[1]	0	99	
14	2	*II*	*ID* indicator	0	10	
15	9	*ID*	identification/call sign	c	c	(Δ •42)
16	2	*C1*	country code	b	b	(Δ •43)
			Regular section (63 characters):			

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
17	1	*DI*	wind direction indic.	0	6	
18	3	*D*	wind direction (true)	1	362	°, 361-2 (Δ dd)
19	1	*WI*	wind speed indicator	0	8	(Δ i_W)
20	3	*W*	wind speed	0	99.9	0.1 m/s (Δ ff)
21	1	*VI*	*VV* indic.	0	2	(Δ •9)
22	2	*VV*	visibility	90	99	(VV)
23	2	*WW*	present weather	0	99	(ww)
24	1	*W1*	past weather	0	9	(W_1)
25	5	*SLP*	sea level pressure	870.0	1074.6	0.1 hPa (Δ PPPP)
26	1	*A*	characteristic of *PPP*	0	8	(a)
27	3	*PPP*	amt. pressure tend.	0	51.0	0.1 hPa (ppp)
28	1	*IT*	indic. for temperatures	0	9	(Δ i_T)
29	4	*AT*	air temperature	–99.9	99.9	0.1°C (Δ s_n, TTT)
30	1	*WBTI*	*WBT* indic.	0	3	(Δ s_w)
31	4	*WBT*	wet-bulb temperature	–99.9	99.9	0.1°C (Δ s_w, $T_bT_bT_b$)
32	1	*DPTI*	*DPT* indic.	0	3	(Δ s_t)
33	4	*DPT*	dew-point temperature	–99.9	99.9	0.1°C (Δ s_t, $T_dT_dT_d$)
34	2	*SI*	*SST* meas. method	0	12	(Δ •30)
35	4	*SST*	sea surface temp.	–99.9	99.9	0.1°C (Δ s_n, $T_wT_wT_w$)
36	1	*N*	total cloud amount	0	9	(N)
37	1	*NH*	lower cloud amount	0	9	(N_h)
38	1	*CL*	low cloud type	0	9, “A”	(Δ C_L)
39	1	*HI*	*H* indic.	0	1	(Δ •9)
40	1	*H*	cloud height	0	9, “A”	(Δ h)
41	1	*CM*	middle cloud type	0	9, “A”	(Δ C_M)
42	1	*CH*	high cloud type	0	9, “A”	(Δ C_H)
43	2	*WD*	wave direction	0	38	
44	2	*WP*	wave period	0	30, 99	seconds (P_WP_W)
45	2	*WH*	wave height	0	99	(H_WH_W)
46	2	*SD*	swell direction	0	38	($d_{W1}d_{W1}$)
47	2	*SP*	swell period	0	30, 99	seconds ($P_{W1}P_{W1}$)
48	2	*SH*	swell height	0	99	($H_{W1}H_{W1}$)

1. Fields differing from the ICOADS-standard representation in the NCDC-variant format (see Supps. D-E for further details). For *MO*, *DY*, and *HR*, the NCDC-variant format uses leading zeros as an exception to the “blank left-fill” aspect of the ICOADS-standard representation for numeric data.

Core update notes:

No format changes, however *IM* is set to “1” in the revised format, to indicate IMMA1. See Annex A for a discussion of NDBC moored buoy wave measurements currently translated into the ship-oriented wave fields *WD*, *WP*, and *WH*. Annex C states that the configuration of *WP* needs to be expanded to include 99, but it already does include 99 here (reason for discrepancy unclear).

Table C1. ICOADS (*Icoads*) attm.

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
1	2	*ATTI*	attm ID			Note: set *ATTI*=1
2	2	*ATTL*	attm length			Note: set *ATTL*=65
			Box elements (6 characters):			
3	1	*BSI*	box system indicator	u	u	(currently set to missing)
4	3	*B10*	10° box number	1	648	(ICOADS BOX10 system)
5	2	*B1*	1° box number	0	99	
			Processing elements (17 characters):			
6	3	*DCK*	Deck	0	999	
7	3	*SID*	source ID	0	999	
8	2	*PT*	platform type	0	15	
9	2	*DUPS*	dup status	0	14	
10	1	*DUPC*	dup check	0	2	
11	1	*TC*	track check	0	1	
12	1	*PB*	pressure bias	0	2	
13	1	*WX*	wave period indicator	1	1	
14	1	*SX*	swell period indicator	1	1	
15	2	*C2*	2nd country code	0	40	
			QC elements (38 characters):			
16-27	1×12	*SQZ-DQA*[1]	adaptive QC flags	1	35	base36 (12 flags)[2]
28	1	*ND*	night/day flag	1	2	
29-34	1×6	*SF-RF*[1]	trimming flags	1	15	base36 (6 flags)[2]
35-48	1×14	*ZNC-TNC*[1]	NCDC-QC flags	1	10	base36 (14 flags)[2]
49	2	*QCE*[3]	external (e.g. MEDS)	0	63	integer encoding (6 flags)
50	1	*LZ*	landlocked flag	1	1	
51	2	*QCZ*[3]	source exclusion flags	0	31	integer encoding (5 flags)

[Note: Detailed QC flag footnotes omitted here.]

Icoads attm update notes:

No changes proposed, other than renumbering of fields.

Table C5. IMMT-5/FM 13 (*Immt*) attm.

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
1	2	*ATTI*	attm ID			Note: set *ATTI*=5
2	2	*ATTL*	attm length			Note: set *ATTL*=94

			Common for IMMT-2/3/4/5 (50 characters):			
3	1	*OS*	observation source	0	6	(•40)
4	1	*OP*	observation platform	0	9	(•41)
5	1	*FM*	FM code version	0	Z	base36 (Δ •64)
6	1	*IMMV*	IMMT version	0	Z	base36
7	1	*IX*	station/weather indic.	1	7	(i_X)
8	1	*W2*	2nd past weather	0	9	(W_2)
	~~1~~	~~*SGN*[1]~~	~~significant cloud amount~~	~~0~~	~~9~~	~~(N_S; ref. Table B3)~~
	~~1~~	~~*SGT*[1]~~	~~significant cloud type~~	~~0~~	~~9, “A”~~	~~(C; ref. Table B3)~~
	~~2~~	~~*SGH*[1]~~	~~significant cloud height~~	~~0~~	~~99~~	~~(h_Sh_S; ref. Table B3)~~
9	1	*WMI*	indic. for wave meas.	0	9	(•31)
10	2	*SD2*	dir. of second. swell	0	38	($d_{W2}d_{W2}$)
11	2	*SP2*	per. of second. swell	0	30, 99	($P_{W2}P_{W2}$)
12	2	*SH2*	ht. of second. swell	0	99	($H_{W2}H_{W2}$)
13	1	*IS*	ice accretion on ship	1	5	(I_s)
14	2	*ES*	thickness of I_s	0	99	cm (E_sE_s)
15	1	*RS*	rate of I_s	0	4	(R_s)
16	1	*IC1*	concentration of sea ice	0	9, “A”	(Δ c_i)
17	1	*IC2*	stage of development	0	9, “A”	(Δ S_i)
18	1	*IC3*	ice of land origin	0	9, “A”	(Δ b_i)
19	1	*IC4*	true bearing ice edge	0	9, “A”	(Δ D_i)
20	1	*IC5*	ice situation/trend	0	9, “A”	(Δ z_i)
21	1	*IR*	indic. for precip. data	0	4	(i_R)
22	3	*RRR*	amount of precip.	0	999	(RRR)
23	1	*TR*	duration of per. *RRR*	1	9	(t_R)
24	1	*NU*	national use	c	c	(national practice)
25	1	*QCI*	quality control indic.	0	9	(•45)
26-45	1×20	*QI1-20*	QC indic. for fields	0	9	(Q_1-Q_{20})
			New for IMMT-2/3/4/5 (41 characters):			
46	1	*QI21*	MQCS version	0	9	(Q_{21})
47	3	*HDG*	ship’s heading	0[2]	360	0, ° (HDG)
48	3	*COG*	course over ground	0	360	0, ° (COG)
49	2	*SOG*	speed over ground	0	99	kt (SOG)
50	2	*SLL*	max.ht.>Sum. load ln.	0	99	m (SLL)
51	3	*SLHH*	dep. load ln.: sea lev.	–99	99	m (s_Lhh)
52	3	*RWD*	relative wind direction	1	362	°, 361-2[3] (ref. *D*)
53	3	*RWS*	relative wind speed	0	99.9	0.1 m/s (ref. *W*)
54-61	1×8	*QI22-QI29*	QC indic. for fields	0	9	(Q_{22}-Q_{29})[4]
62	4	*RH*	relative humidity	0.0	100.0	0.1%
63	1	*RHI*	relative humidity indic.	0	4	(RHi)
64	1	*AWSI*	AWS indicator	0	2	(AWSi)
65	7	*IMONO*	IMO number	0	9999999	(IMOno)

1. Strictly historical fields, moved to *Hist*.
2. Zero is documented to mean “no movement,” but has been suggested should not be used (see Supp. D).

3. Special code 362 for "variable, or all directions" is allocated in IMMA, but IMMT does not presently contain a corresponding configuration for *RWS* (see Supp. D).
4. As from IMMT-4, usage of Q_{26} is discontinued, ref. IMMT-4 documentation: "now Q27 serves as the indicator for both s_L and HH."

Immt attm update notes:

Renumbering of fields.

The current IMMA0 "IMMT-2/FM 13" attm is updated to reflect changes made in three later versions of IMMT: IMMT-3 (effective 1 Jan. 2007), IMMT-4 (1 Jan. 2011), and IMMT-5 (1 June 2012). Differences between IMMT-4 and IMMA0 are documented at http://icoads.noaa.gov/immt4.html. In conjunction with approval by JCOMM-4 of IMMT-5 (http://www.jcomm.info/index.php?option=com_oe&task=viewDocumentRecord&docID=8833), it was decided to discontinue roman, and strictly use (including in WMO publications) arabic, numerals for the format versioning.

A new field (*IMMV*) indicates the applicable IMMT version within the attm, which accommodates some format evolution problems, in that some IMMT fields changed meaning between IMMT-3 and IMMT-4.

For reference, the following table shows selected IMMT versus IMMA field differences outside of this attm, where IMMA generally has enhanced resolution/method:

IMMA *Core* field	IMMT	IMMA
HR	whole hour	0.01 hour
LAT and *LON*	0.1°	0.01° (in *Core*) or 0.001° (in proposed *Auto*)
W	whole kts or whole m/s	0.1 m/s (converted)
ID	7 characters	9 characters

KNMI had suggested that a QC flag for *RH* be considered for addition to IMMT (thus should space be allocated in this revision of IMMA) so that all major elements would have an associated individual QC flag. Also the QC flags (*QCI* and *QI1-20*) are voluminous (occupying 30 characters total) and not optimally organized due to the way IMMT has evolved. If any further changes were made in the *Immt* attm, the option could be considered to bring all the QC flags closer together (including possibly in a separate attm devoted to IMMT QC information).

Table C6. Model quality control (*Mod-qc*) attm. For reference, the Units column also includes (following any units information) the current UK Met Office BUFR element names.

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
1	2	*ATTI*	attm ID			Note: set *ATTI*=6
2	2	*ATTL*	attm length			Note: set *ATTL*=68
			GTS bull. header fields (10 characters):			
3	4	*CCCC*	collecting centre	a	a	COLTN_CNTR
4	6	*BUID*	bulletin ID	b	b	BLTN_IDNY
			Model comp. elements (54 characters):			
5	1	*FBSRC*	Feedback source	u	u	(*tbd*, tentatively base36; e.g. operational vs. reanalysis, or a specific reanalysis)
6	5	*BMP*	background (bckd.) *SLP*	870.0	1074.6	0.1 hPa; BCKD_MSL_PESR
7	4	*BSWU*	bckd. wind U-comp.	–99.9	99.9	0.1 m/s;

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
						BCKD_SRFC_WIND_U
8	4	*SWU*	derived wind U-comp.	–99.9	99.9	0.1 m/s; SRFC_WIND_U
9	4	*BSWV*	bckd. wind V-comp.	–99.9	99.9	0.1 m/s; BCKD_SRFC_WIND_V
10	4	*SWV*	derived wind V-comp.	–99.9	99.9	0.1 m/s; SRFC_WIND_V
11	4	*BSAT*	bckd. air temperature	–99.9	99.9	0.1°C; BCKD_SRFC_AIR_TMPR
12	3	*BSRH*	bckd. relative humidity	0	100	%; BCKD_SRFC_RLTV_HUMDY
13	3	*SRH*	(derived) relative humidity	0	100	%; SRFC_RLTV_HUMDY
	~~1~~	~~*SIX*~~	~~derived stn./wea. indic.~~	~~2~~	~~3~~	~~(subset of *IX*, field 105; unused)~~
14	5	*BSST*	bckd. *SST*	–99.99	99.99	0.01°C; BCKD_SEA_SRFC_TMPR
15	1	*MST*	model surface type	0	9	(UK 008204); MODL_SRFC_TYPE
16	4	*MSH*	model height of surface	-999	9999	m; MODL_SRFC_HGHT
17	4	*BY*	bckd. year	0	9999	year; BCKD_YEAR
18	2	*BM*	bckd. month	1	12	month; BCKD_MNTH
19	2	*BD*	bckd. day	1	31	day; BCKD_DAY
20	2	*BH*	bckd. hour	0	23	hour; BCKD_HOUR
21	2	*BFL*	bckd. forecast length (time period or displacement minute)	0	99	minutes BCKD_FRCT_LNGH

***Mod-qc* attm update notes:**

Renumbering of fields.

Fields *CCCC* and *BUID* are exclusive to operational GTS data. Additional GTS bulletin header information (e.g. fields indicating whether reports are corrected on re-transmission—available from NCEP BUFR at least and now planned for retention as part of the NCDC GTS data) could possibly be useful to also incorporate somewhere in IMMA. However, separating the GTS bulletin header information, from the more specialized model QC fields, might be more efficient, to avoid creating attms that are largely blank. A related consideration is the possibility of composite blending of Met Office BUFR with other GTS sources.

Note: Additionally, and peripherally related to the content of this attm, NCEP has "quality marks" they set, which currently are not preserved in IMMA (ref. http://icoads.noaa.gov/rt.html).

For *BSRH* and *SRH* values appear at least as high as 107%. While actual RH can't be that high, this raises the question whether the ranges of these model-generated fields should be increased in the future e.g. to 107%. On the other hand, *MSH* is extended to a 4-character field, since values of -152.0 and others less than -99 have been detected (plus larger positive values than previously allowed).

Also, *BSST* is translated to SI units at the Met Office using constant 273.15K, whereas a lower-precision 273.1K constant is used for *BSAT*, the only other temperature field presently being made available by the Met Office. To keep its resultant higher precision, *BSST* is expanded to 5 characters. Explanation from Colin Parrett at the Met Office's Real Time Monitoring Centre (RTMC):

> "As far as I know, the conversions depend on the precision of the received data, using 273.0, 273.1 or 273.15 for 0, 1 or 2 (or more) decimal places. I've enquired with our MetDB Team for confirmation and

I'll let you know if things have changed. The background SST does come from a different source, so that might explain the greater precision."

The referenced encoding constant 273.0 does not appear to apply to the temperature elements currently received from the Met Office, but in the event such data were received in the future a 4-character field configuration like that for BSAT would be sufficient (however, to accurately translate temperature data back from Kelvin to °C, it is crucial to know what constant has been used for encoding originally reported °C temperatures to Kelvin for storage in BUFR).

SIX is not reported in the original BUFR files, and there are no plans at the RTMC to begin encoding it—thus it is removed from IMMA1.

Table C7. Ship metadata (*Meta-vos*) attm.

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
1	2	*ATTI*	attm ID			Note: set *ATTI*=7
2	2	*ATTL*	attm length			Note: set *ATTL*=*58*
			Ship metadata elements (54 characters):			
3	1	*MDS*	metadata source	b	b	(base36, with the following defining settings: 0=WMO Pub. 47 vs. 1=COAPS)
4	2	*C1M*	recruiting country	a	a	(Δ •43)
5	2	*OPM*	type of ship (programme)	0	99	(code unlike *OP)*
6	2	*KOV*	kind of vessel	c	c	
7	2	*COR*	country of registry	a	a	(Δ •43)
8	3	*TOB*	type of barometer	c	c	
9	3	*TOT*	type of thermometer	c	c	
10	2	*EOT*	exposure of thermometer	c	c	
11	2	*LOT*	screen location	c	c	
12	1	*TOH*	type of hygrometer	c	c	
13	2	*EOH*	exposure of hygrometer	c	c	
14	3	*SIM*	*SST* meas. method	c	c	(code unlike *SI)*
15	3	*LOV*	length of vessel	0	999	M
16	2	*DOS*	depth of *SST* meas.	0	99	M
17	3	*HOP*	height of visual observation platform	0	999	M
18	3	*HOT*	height of *AT* sensor	0	999	M
19	3	*HOB*	height of barometer	0	999	M
20	3	*HOA*	height of anemometer	0	999	M
21	5	*SMF*	source metadata file	0	99999	e.g. "19991" 1st Q 1991
22	5	*SME*	source meta. element	0	99999	line number in file
23	2	*SMV*	source format version	0	99	~~to be defined~~

***Meta-vos* attm update notes:**

Renumbering of fields.

One new field, *MDS*, is added to indicate the metadata source.

SMV is actually defined, and more information given about other fields specific to WMO Pub. 47, in Dave Berry's documentation regarding the processing of this attm for R2.5 (http://icoads.noaa.gov/e-doc/imma/WMO47IMMA_1966_2007-R2.5.pdf).

Table C8. ICOADS Value-Added Database (*Ivad*) attm.

No.	*Len.*	*Abbr.*	*Element description*	*Min.*	*Max.*	*Units (Code)*
1	2	*ATTI*	attm ID			Note: set *ATTI=8*
2	2	*ATTL*	attm length			Note: set *ATTL=41*
			Value-added data and metadata (33 characters):			
3	2	*ICN*	input component number	0	*(tbd)*	IMMA component number
4	2	*FN*	field number	1	*(tbd)*	IMMA field no. within *ICN*
5	5	*VAD*	value-added data	*(inh.*[1]*)*	*(inh.*[1]*)*	(inherited from *ICN* & *FN*)
6	1	*IVAUR*	indicator for *VAUR*	1	Z	(base36[2])
7	5	*VAUR*	random uncertainty for *VAD*	*(inh.*[1]*)*	*(inh.*[1]*)*	(inherited from *ICN* & *FN*)
8	1	*IVAUB*	indicator for *VAUB*	1	Z	(base36[2])
9	5	*VAUB*	bias uncertainty for *VAD*	*(inh.*[1]*)*	*(inh.*[1]*)*	(inherited from *ICN* & *FN*)
10	1	*IVAUA*	indicator for *VAUA*	1	Z	(base36[2])
11	5	*VAUA*	additional uncertainty for *VAD*	*(inh.*[1]*)*	*(inh.*[1]*)*	(inherited from *ICN* & *FN*)
12	2	*VQC*	value-added QC flag	0	9	See notes and Table C8a
13	2	*ARC*	author reference code–*Ivad*	b	b	[Note: alphanumeric, thus also allows interpretation as base36]
14	2	*AJDN*	archive adjusted Julian day number[3]	0	ZZ	(base36)

1. The range and other characteristics of these value-added data fields are all inherited from *ICN* & *FN*.
2. Indicator configurations to be fully defined during prototyping.
3. Intended to document insertion of this attm into the ICOADS IMMA archive. Julian day number is the integer part of the Julian date (ref. http://en.wikipedia.org/wiki/Julian_day#Julian_Date), but tentatively *AJDN* will instead start with day zero at 1 October 2012 (thus *AJDN* + 2456201 would equal the official Julian Day number). ZZ is 46655 in base36, thus allowing ~3.5 years of entries (in contrast using 3-character, ZZZ is 46655 in base36, thus use of three characters would allow almost 128 years of entries).

IVAD attm notes (new attm)

See Annex D for a general discussion of scenarios for adding adjustments or QC within the IVAD project, and see Annex E for a more specific discussion of *Ivad* attm field configuration and record management details (with the UK National Oceanography Centre).

Bias adjusted fields will be stored in this *Ivad* attm, whereas the unadjusted data will be stored in the *Core*/other attms. Note that this is an inversion of the planned handling, after blending into ICOADS, of straightforward data corrections using the *Error* attm (see Table C9).

Rather than including these attms (of which there may be a varying number) within the main IMMA record, we are implementing a Subsidiary *Ivad* record-type to store an indefinite number of individual *Ivad* attms together in separate physical record(s), for a given *UID*. This requires that both the Main and Subsidiary record types include *UID* (together with ICOADS Release number details *RN1*+*RN2*+*RN3*) so that they can be identified (i.e. by Subsidiary record type) and linked back together with the main record. This linking is accomplished using the new *Uid* attm (see Table C98):

Main IMMA record type: *Core* + *Icoads* + *Immt* + *Mod-qc* + *Meta-vos* + *Uid* + *Suppl*
Subsidiary IMMA record type: *Uid* + *Ivad* + *Ivad* + *Ivad* … + *Ivad*
Subsidiary IMMA record type: *Uid* + *Error* + *Error* + *Error* … + *Error*
Subsidiary IMMA record type: *Uid* + *Meta-vos*

Unresolved questions concerning this new multi-record (Main and Subsidiary) linked approach include: whether there can be multiple subsidiary records associated with a given main record *UID* (probably the most flexible approach), and what limit may need to be established on the overall number of *Ivad* and/or *Error* attms associated with a given main record UID (for Fortran memory management considerations; 100 is tentatively suggested as the maximum).

For *VQC*, we envision this as a mechanism for storing externally provided data QC information, such that the provider of QC information would be required to map their flags to the *VQC* configuration (Table C8a) and describe their mapping method in external documentation as linked via *ARC* (also original flags could be stored in the *Suppl* attm together with original data).

Table C8a. Proposed configuration of the value-added QC Flag (VQC)[1]. Annexes E and F provides additional background, including a possible alternative scheme (Table E2).

Flag	*Data quality indicators*
0	No QC has been performed on this element
1	QC has been performed; element appears to be correct
2	QC has been performed; element appears to be inconsistent with other elements
3	QC has been performed; element appears to be doubtful
4	QC has been performed; element appears to be erroneous
5	The value has been changed as a result of QC
6	Reserved
7	Reserved
8	Reserved
9	The value of the element is missing

1. Adapted partially from the Minimum Quality Control Standard (MQCS) configuration for flags Q_1-Q_{29} as stored in the IMMT format (ref., http://www.bom.gov.au/jcomm/vos/documents/immt4.pdf). However MQCS flags 6 and 7:

6 – The flag as received by the GCCs was set to “1” (correct), but the element was judged by their MQCS as either inconsistent, dubious, erroneous, or missing

7 – The flag as received by the GCCs was set to “5” (amended) but the element was judged by their MQCS as inconsistent, dubious, erroneous, or missing

are specialized to Global Collection Centre (GCC) functions and not widely meaningful outside that environment. Thus they were not used for *VQC*, and are marked as “Reserved” (e.g. for IVAD-related QC).

Table C9. Error (*Error*) attm.

No.	*Len.*	*Abbr.*	*Element description*	*Min.*	*Max.*	*Units (Code)*
1	2	*ATTI*	attm ID			Note: set *ATTI*=9
2	2	*ATTL*	attm length			Note: set *ATTL=(inh.*[1]*)*
			Corrected erroneous data and metadata:			
3	2	*ICN*	input component number	0	*(tbd)*	IMMA component number
4	2	*FN*	field number	1	*(tbd)*	IMMA field no. within *ICN*
5	1	*CEF*	corrected/erroneous field flag	0	1	0: *ERRD* is the corrected value; 1: *ERRD* is the erroneous value
6	*(inh.*[1]*)*	*ERRD*	corrected/erroneous field value	*(inh.*[1]*)*	*(inh.*[1]*)*	(inherited from *ICN* & *FN*)
7	2	*ARCE*	author reference code–*Error*	b	b	[Note: alphanumeric, thus also allows interpretation as base36]
8	2	*AJDNE*	archive adjusted Julian day number—*Error*	0	ZZ	(base36; as for *AJDN*, ref. Table C8)

Error attm notes (new attm)

Implementation will be handled similarly to the *Ivad* attm (Table C8) in that any number of individual *Error* attms will be stored in one (or more) *Error* Subsidiary record(s), and linked together using the *Uid* attm (Table C98). *CEF* distinguishes between *Error* Subsidiary records as provided externally to ICOADS (*CEF*=0), in contrast to after the *Error* records are blended into ICOADS (*CEF*=1).

To simplify the user interface, the plan is that corrections for straightforward errors (e.g. callsign garbling) will ultimately be stored by ICOADS in the *Core*/other attms, whereas uncorrected data will be stored in this *Error* attm—this is an inversion of the planned handling of bias adjustments using the *Ivad* attm. The swapping of the information in the provided *Error* attms, to final inverted storage in IMMA (i.e. from *CEF*=0 to *CEF*=1, and interchanging the data fields), will probably be handled centrally; however, the *CEF* flag settings should allow this inversion to be handled externally instead if desired (i.e. through the provision of both Main and Subsidiary records).

Although some reduction in data volume is achieved by having the wide of the data field vary (i.e. inherited from the referenced field configuration), this remains a relatively voluminous approach to storing error information for individual fields owing to the attm overhead. Alternatively, similar to the old LMR "error attachment" (ref. discussion in red at the end of http://icoads.noaa.gov/Release_1/suppF.html regarding "Attachment 5") this could potentially be a variable-length attm of the following form: *FN*, *ERRD*, *NREP* number of repetitions (i.e. 1/more repetitions of *FN*+*ERRD*, …, all falling under one *ARCE*).

Table C98. Unique ID (*Uid*) attm.

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
1	2	*ATTI*	attm ID			Note: set *ATTI*=98
2	2	*ATTL*	attm length			Note: set *ATTL*=15
			Processing elements (10 characters):			
3	6	*UID*	Unique report ID	b	b	base36[1]
4	1	*RN1*	Release no.: primary	0	Z	base36, e.g. 2
5	1	*RN2*	Release no.: secondary	0	Z	base36, e.g. 5
6	1	*RN3*	Release no.: tertiary	0	Z	base36, e.g. 0 (thus 2.5.0 together)
7	1	*RSA*	Release status indicator	0	2	0=Prelim., 1=Aux., 2=Full
8	1	*IRF*	intermediate reject flag	0	1	0=Reject, 1=Retained

1. While it represents a base36 number, this field is handled by rdimma1 as strictly (i.e. without leading spaces, e.g. 35=00000Z) alphanumeric, and thus is not fully translated into an integer or floating-point (REAL) number (ref. rdimma1 comments: "For character […] fields, note that ITRUE and FTRUE contain the ICHAR of the first character of the field…"). Users interested in handling *UID* as a number should be aware of possible finite precision issues arising in the representation of large numbers on computers:

- In the integer case, the largest 6-character base36 number is ZZZZZZ (2,176,782,335); however, if one bit is reserved for sign, the largest positive integer representable in 32 bits is only 2^{31}–1 (2,147,483,647; ZIK0ZJ in base36). However as noted below the current maximum of UID is $m_{R2.5i}$ (~295M) and thus well below this threshold.
- Whereas, in the floating-point case it is not even possible to accurately represent $m_{R2.5i}$ as a 32-bit single precision REAL number.

Uid attm notes (new attm):

The *intermediate* Release 2.5 product (R2.5*i*), containing available duplicates and other reports excluded from the normal user product (R2.5), was used as the starting point for assigning *UID* and these other new fields, and generally prototyping IMMA1. The Release number fields were set to indicate R2.5.1 (i.e. *RN1*=2, *RN2*=5, *RN3*=1), and *IRF* to indicate whether each report is to be rejected or retained during construction of a final user product R2.5.1, from the prototype intermediate product R2.5.1*i* (note: both the intermediate R2.5.1*i* and the final R2.5.1 data are available to interested users).

R2.5*i* contains ~295M (specifically: 294,725,525) reports ($m_{R2.5i}$), so all those records (in predefined temporal archive sequence) had *UID* assigned from 1,…,$m_{R2.5i}$. In preparation for processing the next Release, we anticipate that all new and historical records will be numbered starting from $m_{R2.5i}$+1. After blending the old and new records into the new Release, all the *UIDs* will no longer be sequential (i.e. new *UIDs* will be interleaved into the old purely numeric sequence; see Annex B for further discussion).

One additional idea that came up in ETMC-III (2010) discussion was to allow for the possibility as well for a national *UID*, e.g. *UIDN*, as opposed to the main international *UID*. This should be kept under consideration as a possibility for the future, but we do not recommend expanding e.g. *Uid* with that information (which would seem to fit better into some other attm).

Table C99. Supplemental data (*Suppl*) attm. If *ATTL*=0 (unspecified length), this attm must appear at the end of the record, and the record terminate with a line feed. For the VOSClim record type, this attm stores the original input data string in Ascii with *ATTL*=0 and *ATTE*=missing. (Note: if future requirements arise within the VOSClim record type, or for other record types, *ATTL* and *ATTE* can be adjusted accordingly.)

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
1	2	*ATTI*	attm ID			Note: set *ATTI*=99
2	2	*ATTL*	attm length			Note: set *ATTL*=0
3	1	*ATTE*	attm encoding			Note: set *ATTE*=missing
			Supplemental data (format determined by data source):			
4		*SUPD*[1]	supplemental data	c	c	

1. The length of the supplemental data is *ATTL* – 5 if *ATTL* > 0, or it may be variable if *ATTL* = 0.

Suppl attm update notes:

No changes proposed, other than renumbering of fields, and in the table numbering to match *ATTI*.

Table CP1. Physical oceanographic (*Pocn*) attm.

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
1	2	*ATTI*	attm ID			Note: set *ATTI=(tbd)*
2	2	*ATTL*	attm length			Note: set *ATTL=45*
			Oceanographic data and metadata (41 Characters):			
3	5	*OTV*	temperature value	*-3.*	*35.0*	0.001°C

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
4	1	*OTP*	original temperature value precision			e.g. 2 = 0.01°C precision
5	5	*OTZ*	temperature depth	*0.*	*10.*	0.001 m [max. set to agree with WOD05 transpec procedures]
6	1	*OTZP*	original temperature depth precision	*0*	*2*	[numerals right of decimal]
7	5	*OSV*	salinity value	*0.*	*40.*	0.001 [unitless, on Practical Salinity Scale]
8	1	*OSP*	original salinity value precision	*0*	2	[numerals right of decimal]
9	5	*OSZ*	salinity depth	*0.*	*10.*	0.001 m [max. set to agree with WOD05 transpec procedures]
10	1	*OSZP*	original salinity depth precision	*0*	*2*	[numerals right of decimal]
11	2	*OPT*	probe Type, 2nd header 29, with caveat to reset…	*0*	*16*	WOD09 s_29_probe_type [WOD code]
12	1	*RST*	reference scale for temperature	*102*	*103*	WOD09 v_3_scale [WOD code]
13	1	*RSS*	reference scale for salinity	*202*	*203*	WOD09 v_3_scale [WOD code]
14	2	*NSB*	normal standard seawater batch number	*NA*	*NA*	IAPSO batch number, Variable specific Secondary Header (18) [IAPSO batch number]
15	1	*OIC*	instrument calibration	*NA*	*NA*	1=uncalibrated, Variable specific Secondary Header (16) [WOD code]
16	5	*PAT*	Argo profiler adjustment value for temperature	*-2.*	*2.*	0.001°C
17	5	*PAS*	Argo profiler adjustment value for salinity	*-0.5*	*0.5*	0.001 [unitless, on Practical Salinity Scale]

***Pocn* attm notes:**

Field contents tailored to the specialized requirements of capturing near-surface data deemed most relevant to marine meteorology from the World Ocean Database (e.g. WOD09; http://www.nodc.noaa.gov/OC5/WOD/pr_wod.html).

Unresolved questions (E. Freeman) regarding listed range [0,2] of *OSP*: "Will this work? These appear to be unit-less in WOD and according to '9.2 Salinity' in the WOD documentation (pp. 147-149) they are reported to hundredths with hundredths value always '0'. With new and higher precision instruments, a little cushion to thousandths might not be a bad idea."

Table CP2. Automated instrumentation (*Auto*) attm (proposed)

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
1	2	*ATTI*	attm ID			Note: set *ATTI=(tbd)*
2	2	*ATTL*	attm length			Note: set *ATTL=41*
			Automated instrumental metadata (37 characters):			
3	8	*ALAT*	latitude	–90.000	90.000	0.001°N

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
4	9	*ALON*	longitude	0.000	359.999	0.001°E (ICOADS conv.)
5	1	*INAV*	navigation system indicator	0	9	(controlled vocabulary *tbd*, e.g. 0=GPS, 1=POSMV, 2=INS)
6	6	*APRS*	atmospheric pressure	870.00	1074.60	at barometer height (*HOB*)
7	6	*ARSW*	shortwave radiation	0.00	1600.00	Wm^{-2}
8	1	*IARSW*	shortwave radiation indicator	0	9	(controlled vocabulary *tbd*, e.g. 0=down-, 1=upwelling)
9	5	*ARLW*	longwave radiation	200.00?	800.00?	Wm^{-2}
10	1	*IARLW*	longwave radiation indicator	0	9	(controlled vocabulary *tbd*, e.g. 0=down-, 1=upwelling)

***Auto* attm notes:**

This attm was designed to provide a location to capture meteorological and underway ocean data that are not routinely reported by VOS or in historical ship reports. These values would be derived from automated instrumentation.

This attm could be expanded to include all possible parameters that could be derived at high precision from automated instrumentation. Candidate fields that are included elsewhere in IMMA0 are: Ship's course and speed (*DS*/*VS*, in the *Core*; or *COG/SOG* for the over ground elements, in *Immt*), and ship's heading (*HDG* in *Immt*), wind direction and speed (true *D*/*W*, in the *Core*; or relative *RWD/RWS* in *Immt*), AT, WBT, DPT (*Core*), and RH and precipitation (*Immt*). Other possible fields for this table include visibility and cloud height derived from automated sensors, but they are currently very rare on ships or moorings, or possibly surface velocity data (not presently part of ICOADS).

For *ARSW*, we need to decide if we want to allow for negative values. They are common due to sensor calibration issues (and flagged e.g. by SAMOS), but are not physical.

Storing *APRS* is proposed for two reasons (a) there is no place in IMMA to store atmospheric pressure values not converted to sea level and (b) precision automated barometers can easily record *SLP* (or *APRS*) to 2 or 3 decimal places. However, if the field serves two purposes, an associated indicator may be needed to flag the high-resolution pressure type (i.e. *SLP* or *APRS*)

Radiation could be handled in different ways. The idea above provides for separate shortwave/longwave total radiation variables. If we added a signed range, this could also allow for net radiation. Another other option would allow for multiple radiation values each with an indicator stating whether it is shortwave, longwave, PAR, UV, etc. This may result in a variable-length attachment or one of fixed-length with many empty fields. Also, some indicator of the time period over which the radiation was integrated may be needed. The draft E-SURFMAR Dataformat#100 (http://esurfmar.meteo.fr/doc/o/vos/E-SURFMAR_VOS_formats_v011.pdf) suggests "over the past hour." More discussion is needed on these issues.

Table CP3. Platform tracking (*Track*) attm (proposed).

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
1	2	*ATTI*	attm ID			Note: set *ATTI=(tbd)*
2	2	*ATTL*	attm length			Note: set *ATTL=?*
			Platform track information (xx characters):			
3	1?	*UIDT*	UID type			(*tbd*; e.g. 1=ICOADS-

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
						standard, 2=collection/*SID*-specific,3=platform/voyage-specific)
4	6	*UID1*	UID of previous report	1	*(tbd)*	
5	6	*UID2*	UID of this report	1	*(tbd)*	
6	6	*UID3*	UID of next report	0	*(tbd)*	
7	2	*ARCT*	author reference code–*Track*	b	b	[Note: alphanumeric, thus also allows interpretation as base36]
8	2	*AJDNT*	archive adjusted Julian day number—*Track*	0	ZZ	(base36; as for *AJDN*, ref. Table C8)

***Track* attm notes:**

Sets aside space for "pointer" fields indicating the UID of the previous (*UID1*) and next (*UID3*) report, with respect to this report (*UID2*), in ship/buoy track sequence (i.e. both forward, and backward, in time and space). If indicated by *UIDT*, this attm could contain collection- (or source ID, *SID*) specific, or even platform/voyage-specific, rather than ICOADS-standard, *UID* information (which thus in a sense can be considered value-added information, if assembled externally).

This could be very useful e.g. for reanalyses to resolve the problem of connecting ship/buoy voyages within ICOADS. Due to effects of dupelim, tracks may consist of records interspersed from a variety of sources, with possibly varying *IDs* for records in track sequence. This proposed attm would provide the storage mechanism for this information, but populating the attm seems likely to be challenging and is not presently resourced. Therefore as with the *Ivad* attm, we might consider this to be metadata we would consider ingesting if somebody else had the resources to implement the ship tracking. [Note: Meanwhile a related improvement would be to try to fill in more ship callsigns etc., e.g. through substitution of information among duplicates.]

Table CP4. Historical attm *(Hist)* (proposed). *ATTI* to be assigned, and *ATTL* and field numbering to be decided (*tbd*).

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
1	2	*ATTI*	attm ID			Note: set *ATTI=(tbd)*
2	2	*ATTL*	attm length			Note: set *ATTL=(tbd)*
			Historical data fields (>19 characters):			
3	?	*SN*	ship's name	u	u	[Note: either the full name, or possibly abbreviated with reference to a separately maintained list, to same space?]
4	5	*LCR*	longitude by chronometer	0.00	359.99	0.01°E[1] (ICOADS conv.)
5	5	*LMG*	longitude made good[2]	0.00	359.99	0.01°E[1] (ICOADS conv.)
6	5	*LDR*	longitude by account[3]	0.00	359.99	0.01°E[1] (ICOADS conv.)
7	1	*WFI*	*WF* indic.	u	u	
8	2	*WF*	wind force	0	12	
9	1	*XWI*	*XW* indic.	u	u	
10	3	*XW*	wind speed (ext. *W*)	0	99.9	0.1 m/s
11	1	*XDI*	*XD* indic.	u	u	
12	2	*XD*	wind dir. (ext. *D*)	u	u	

No.	Len.	Abbr.	Element description	Min.	Max.	Units (Code)
14	1	*SLPI*	*SLP* indic.	u	u	[Note: This or another indicator needed to indicate the presence or absence of *SLP* adjustment (ref. *PB*)?]
15	1	*TAI*	*TA* indic.	u	u	
16	4	*TA*	*SLP* att. thermometer	–99.9	99.9	ref. *AT*
17	5	*SMPR*	sympiesometric pressure	25.000	32.000	0.001 inches of mercury[4]
18	1	*XNI*	*XN* indic.	u	u	
19	2	*XN*	cloud amt. (ext. *N*)	u	u	
20	1	*SGN*	significant cloud amount	0	9	(N_S; ref. Table B3)
21	1	*SGT*	significant cloud type	0	9, "A"	(C; ref. Table B3)
22	2	*SGH*	significant cloud height	0	99	($h_S h_S$; ref. Table B3)
(plus additional elements tbd)						

1. A possible alternative approach for storing these longitudes, such as from the EEIC collection, would be to keep the DDD.MM.SS original format, noting however that original data configurations should be preserved anyway in the *Suppl* attm. Also storing decimal points would violate the standard IMMA representation for numeric data (unless these fields were stored as character strings).
2. With reference to Greenwich Meridian.
3. As calculated by dead reckoning.
4. Due to the erratic nature of the sympiesometer measurements such as observed in the EIC Collection, these values might fall well out of the range specified here.

Hist attm update notes:

Fields *SGN*, *SGT*, and *SGH*, which are believed to be purely historical (1960s or earlier), are moved here from the *Immt* attm. Refer to the complete version of the IMMA0 documentation for Table B3 (http://icoads.noaa.gov/e-doc/imma/R2.5-imma.pdf). Among potential additional elements: dead reckoning positions (if preserved additionally to observed positions) and surface current movement (derivable from dead reckoning positions), Leeway, magnetic deviation and variation, etc.

Other examples from recent work on the C19th German Maury Collection:

Cloud form:

Cirrus	CI	Cirrocumulus	CC	Cirrostratus	CS
Altocumulus	AC	Altostratus	AS		
Stratocumulus	SC	Stratus	ST	Nimbostratus	NS
Cumulus	CU	Cumulonimbus	CB		

Present Weather indicated by combinations of the following Beaufort Codes:

b	blue sky	p	passing showers
c	cloudy sky	q	squally
d	drizzle	r	rain, rainy
f	fog	s	snow
g	gloomy	t	thunder
h	hail	u	ugly threatening sky
l	lightning	v	exceptional visibility
m	mist	w	dew
o	overcast, overcast skies	z	haze

Additional historical fields, such as the following, will have to be investigated much further to determine the feasibility of incorporating them in IMMA. Historically, these are largely non-standardized recordings, recorded in comments possibly embedded in large amounts of text (e.g. >1500 unique state of sea and weather comments in the EEIC collection).

Historic sea state
Historic sea ice

Annex A: Implementation/Unresolved Issues

(a) Handling higher-resolution SST

New requirements to store higher resolution SST data (e.g. to hundredths of a degree) are emerging e.g. from buoys and C-MAN. Different options to handle these data could include: (i) extending the precision of SST by one decimal place in the *Core* (plus probably requiring additional indicator settings); (ii) storing the data together with subsurface ocean temperatures within the *Pocn* attm; (iii) through an expansion of the *Auto* attachment.

In the *Core*, *IT* (indicator for temperatures) could hold the precision information, but since it refers to all temperature values in a single report (multiple temperatures in a single record could be to different precisions), this raises the question whether a precision indicator for each individual temperature field (*SST*, *AT*, *WBT*, and *DPT*) could be useful. Past OISST work at NCDC has encountered problems using the existing *IT* information, but because it can be ambiguous the information has not been considered generally very useful.

(b) Questions about mixing ship and buoy wave data (or other fields)

Some NDBC wave data currently are transformed for storage in Table C2 fields (potentially inappropriate). Specifically, ICOADS contains increasing amounts of measured wave data from NOAA National Data Buoy Center (NDBC) moored buoys in the vicinity of the US coastline. Specifically, these variables in the NCDC TD-1171 format (NCDC 2003) have been translated into IMMA variables (with a loss of data resolution, at least in the case of *WD*, which is represented in degrees in TD-1171 (e.g. 0-360) as compared to coded units of ten degrees in IMMA, e.g. 0-36):

WD = principal wave direction (pos. 84-86)
WH = significant wave height (pos. 75-77)
WP = dominant wave period (pos. 78-80)

In the future, ICOADS should consider adding additional sources of measured wave data, e.g. Canadian and E-SURFMAR.

(c) Mod-qc alternative implementation ideas (probably for envisioned Rean attm)

Possibly this could be generalized, or another attm could be developed (leaving this one as is for the near term to handle the more specialized existing VOSClim requirements) to handle model QC/feedback information akin to the suggested generalize IVAD approach (i.e. keyed to *FN*). This idea hinges on the suggestion that reanalysis feedback information would probably be *FN*-specific, and has fields for background value, estimated value, and then a reference indicator (e.g. UK VOSCLIM, ERA-40, etc.) to cover operational and reanalysis model feedback.).

(d) Gathering Requirements for Future VOS Delayed Data Format (from 2010)

Future/additional requirements (ref. e-mail 2 April 2010 from Frits Koek and Martin Stam)

"It is very difficult to think about the future usage of VOS data. Nevertheless we gave it a go.
Extra elements we could think of and would like to add are:

- Wind speed at anemometer height;
- Wind speed reduced to 10m;
- Anemometer height;
- Method of reduction of the wind speed to 10m;
- Depth of SST measurement below water level;

- Air pressure at instrument level;
- Air pressure at sea level;
- Height of barometer above sea level;
- Correction method for reduction to sea level.

Further, in general, it would be ideal if per element at least the following information is available:

- Actual measurement/reading;
- Method (instrument/visual/calculated);
- Precision (thousandths, hundredths, tenths, whole, etc.);
- Sample size (1s, 1min, 5min, 10min, etc.);
- Average/median/instantaneous;
- Make and model of instrument;
- Location of instrument;
- Exposure of instrument;
- Height of instrument above sea level;
- Units."

Eric Freeman 8 April 2010 e-mail: In addition to the elements that Frits and Martin have recommended, here are a few more that may be handy in the future:

- One major topic that has been mentioned a few times is the removal of the quadrant field and the addition of positions in the latitude and longitude fields to account for higher precision as well as a minus (-) sign for the hemisphere.
- Salinity
- Depth of sfc salinity reading.
- Radiation (sw & lw)
- Current (sfc or near sfc)
- Depth of current reading

Many of these should help with flux calculations and satellite calibration/validation. There may also need to be additional metadata/qc fields associated with these.

Annex B: *Planned ICOADS Development of a Unique Report ID (UID); and Intra-record Release No. (RN) Tracking* [note: adapted from material provided in 2010 to the Data Provenance working group of the Surface Temperature Initiative]

IMMA format improvements in the context of the IVAD project will include a Unique Record ID (UID), in addition to improved ICOADS Release number (RN) tracking (i.e. within the IMMA records, as is not presently the case). The specific implementation of the UID will be coordinated e.g. with reanalysis centers during the development (and prototyping) process to solicit any additional ideas or technical considerations. The prospect has already been raised whether other developing or planned in situ international "comprehensive" archives (e.g. upper-air and land surface) should considering compatible (as applicable) or at least technically similar schemes (e.g. so the UID interface with reanalyses to the various archives could be unified).

Initial assignment of the UID number to the archive
In the simplest form, the UID might be initiated for ICOADS starting with a numbering from 1, …, ~295M ($m_{R2.5}$) of all the records (in some precise sequential archive ordering to be defined, but probably temporally organized at the highest sort level, e.g. 1662, 1663, …) in the R2.5 intermediate product (ICOADS, 2010) (however the possible advantages of having the UID be a smaller number instead local to an archive increment, such as year-month, should also be considered). The intermediate product, described in ICOADS (2010), contains all currently blended duplicates and other questionable reports, and from it a smaller finalized product (261M reports) without the dups etc. is constructed for most users (although the intermediate product is also available in the event advanced users wish to study the duplicate matching etc.).

Tentative plan for handling new records introduced during Releases/updates
Number these from $m_{R2.5}$ + 1 to $m_{R2.k}$ (k = new Release increment, for an additional set of records to be blended). When these data are blended into e.g. R2.5 data (i.e. records numbered 1, …, ~295M ($m_{R2.5}$), in the resulting blended data the UIDs will no longer be sequential (i.e. new UIDs will be interleaved into the old purely numeric sequence).

Possibility of merged (multi-source) reports in the future
Data sources such as the VOS Climate (VOSClim) project, in which unique data fields flow from up to three distinct data streams (GTS, delayed-mode, and NWP comparison) could benefit from merged records in the future. However, blending records could potentially be used more widely to improve the quality and completeness of the data more generally (e.g. GTS vs. delayed-mode). Probably blended records should receive a new UID (and DCK, SID).

UID format considerations (database and IMMA)
Base36 (alphanumeric) encoding (ICOADS 2010, Table 1) could potentially be used (both in the DBMS and IMMA, if desired, or the DBMS could decode the base36 values into integers). This could achieve at least a 50% reduction (?) in storage size. Dave Berry (UK NOCS) has implemented generalized software for this purpose, and different implementations appear possible either to maximize space savings, at the expense of CPU time, or vice versa.

Annex C: Reprocessing Notes for Recovering Missing Field Configurations, Etc.

Following are excerpts from the current IMMA0 documentation (http://icoads.noaa.gov/e-doc/imma/R2.5-imma_short.pdf) for individual fields discussing proposed, or potential additional, field configuration modifications. For expanded field configurations (e.g. *WP*=99 currently missing in IMMA), this raises the question whether data (e.g. attached supplemental data, if available) should be reprocessed as resources permit to populate the missing configurations (generally these aspects are highlighted in yellow below).

In addition, we note that newly defined fields in the revised *Immt*, *Mod-qc*, and *Meta-vos* attms generally will not be populated in the prototype IMMA (intermediate) dataset (R2.5.1*i* and R2.5.1), with the exception of field MDS in the *Meta-vos* attm.

As an important related issue however, a translation software update to IMMA1 from IMMA0 will be desirable, as resources permit (and probably subsequent to the prototyping phase, to ensure the finality of IMMA1 details), for newly available historical data sources, as well as operational contemporary data sources (e.g. GTS datastreams, and IMMT/IMMA data flowing regularly from the GCCs).

30) *WBTI* *WBT* indicator

32) *DPTI* *DPT* indicator […]

Background: *WBTI* and *DPTI* are derived from sign positions s_w and s_t in IMMT-4. [Note: For data originally translated into LMR from IMMT formats, the predecessor LMR field *T2* preserved only a subset of information derived from s_w and s_t, coupled with whether *DPT* was computed during ICOADS processing. Future work should seek to recover more complete information for data that were translated to IMMA from LMR, and consider new configurations to separately document ICOADS processing. WMO (2009a) Reg. 12.2.3.3.1 specifies when (e.g. owing to instrument failure) relative humidity (RH) is available and may be reported in FM 13 instead of *DPT* in an alternative group 29UUU. Thus far such RH data have generally not been recovered into ICOADS).]

36) *N* total cloud amount (cover)

37) *NH* lower cloud amount […]

38) *CL* low cloud type […]

39) *HI* cloud height indicator […]

40) *H* cloud height […]

41) *CM* middle cloud type […]

42) *CH* high cloud type […]

Background: Configurations for *CL*, *H*, *CM*, and *CH* are as in IMMT-4, except for use of "A" (10 in base36) in place of "/" (LMR used 10 in place of "/"). Analyses of cloud types may be impacted by a 1 Jan. 1982 GTS code change: When *N*=0, the types *CM*, *CH*, and *CL* were reported as missing (i.e. the FM 13 $8NhC_LC_MC_H$ group was omitted), whereas previously these types may have been reported zero (see Hahn et al. 1992). However, to improve climatological data quality, starting 2 Nov. 1994 FM 13 was again modified so that all cloud observations at sea including no cloud observation shall be reported (see WMO 2009a, Reg. 12.2.7.1). [Note: For historical reasons (see background under *NH*, field 37), an inconsistency exists in IMMA in how solidus ("/") is translated for *N* and *NH* (i.e. to missing data) versus for *CL*, *H*, *CM*, and *CH* (i.e. to "A"). A related complication (i.e. in terms of preserving information about whether data were explicitly reported as "/" versus omitted from transmission) is that group Nddff in FM 13 is mandatory, whereas $8NhC_LC_MC_H$ can be omitted (Reg. 12.2.7.1).]

44) *WP* wave period […]

Background: Historically, the (wind) wave and swell codes have been subject to complex changes. Prior to 1949 both sets of fields were apparently reported descriptively in the SHIP code, and thus are expected to be missing (and the swell fields are expected to be missing prior to 1 July 1963, as discussed below). Codes 37-38 arise from earlier historical codes (see Met Office 1948). Starting in 1968, *WD* was no longer reported and *WP* was reported in seconds. [Note: *WP*=99, indicating a confused sea, is not presently defined in IMMA. Future work should seek to recover this information from original formats, and consider an expanded IMMA configuration.]

Discussion: The current configuration of *WP* in IMMA0 includes *WP*=99, so the above highlighted note may no longer fully apply (further investigation needed).

105) *IX* station/weather indicator […]

Background: Starting 1 Jan. 1982, the procedure for reporting present (*WW*) and past (*W1*, *W2*) weather in FM 13 was altered significantly by adding *IX*, which allowed the "7 group" ($7wwW_1W_2$ for manual stations, and usually $7w_aw_aWa_1Wa_2$ for automatic stations) to be omitted when there was no significant present or past weather to report (see Hahn et al. 1992). However, to improve climatological data quality, starting 2 Nov. 1994 FM 13 was again modified so that any present and past weather including phenomena without significance shall be reported (see WMO 2009a, Reg. 12.2.6.2). [Note: Refer to the LMR documentation for more information regarding use of *IX* with present and past weather data, and unforeseen complications attending its introduction in 1982 (e.g. *IX* was not included in IMMT until 1 March 1985). *IX*=4 was initially defined (WMO 1981) without the Code references (hence brackets above), and *IX*=7 was introduced at a later date. The *IX*=7 value was not included in LMR, thus future work should seek to recover this information for data that were translated to IMMA from LMR.]

107) *SGN* significant cloud amount

108) *SGT* significant cloud type

109) *SGH* significant cloud height […]

Background: These significant cloud fields are listed in Met Office (1948), but appear to have been omitted from regular IMM fields (see Table B3) and the current FM 13 code; in presently available ICOADS data they should always be missing [Note: Since these appear to be strictly historical fields, deletion from this attachment and possible repositioning within Table C5 is suggested for future consideration).]

Discussion: In the proposed revision, these fields are moved into the *Hist* attm.

110) *WMI* indicator for wave measurement […]

Background: Note: Field not included in the LMR regular section, thus future work should seek to recover this information for data that were translated into IMMA from LMR.]

111) *SD2* swell direction (2nd)

112) *SP2* swell period (2nd)

113) *SH2* swell height (2nd) […]

Background: [Note: Fields not included in the LMR regular section, thus future work should seek to recover this information for data that were translated into IMMA from LMR.]

117) *IC1* concentration of sea ice

118) *IC2* stage of development
119) *IC3* ice of land origin
120) *IC4* true bearing ice edge
121) *IC5* ice situation/trend […]

Background: Separate fields (or an Code indicator) could be considered in the future. Earlier historical ice codes might also need to be researched for possible consideration. Met Office (1948) lists an Ice Group (c_2KD_ire) that may be similar or identical to the above pre-1982 code (see also Table B3 of Supp. B). [Note: Fields not included in the LMR regular section, thus future work should seek to recover this information for data that were translated into IMMA from LMR.]

122) *IR* indicator for precipitation data
123) *RRR* amount of precipitation
124) *TR* duration of period of reference for amount of precipitation […]

Background: [Note: Fields not included in the LMR regular section, thus future work should seek to recover this information for data that were translated into IMMA from LMR.]

125) *QCI* quality control (QC) indicator […]

Background: Prior to IMMT-4, values 7-8 were instead termed "not used." [Note: Field not included in the LMR regular section, thus future work should seek to recover this information for data that were translated into IMMA from LMR.]

126) *QI1* QC indicator for height of clouds […]
145) *QI20* QC indicator for ship's position […]

Background: [Note: Fields not included in the LMR regular section, thus future work should seek to recover this information for data that were translated into IMMA from LMR, plus additional QC indicators available in IMMT-3/-4.]

146) *QI21* MQCS version […]

Background: [Note: Field not included in the LMR regular section, thus future work should seek to recover this information for data that were translated into IMMA from LMR.]

147) *HDG* ship's heading […]
148) *COG* course over ground […]
149) *SOG* speed over ground […]
150) *SLL* max.ht.>Sum. load ln. […]
151) *SLHH* departure of Summer max. load line from actual sea level […]
152) *RWD* relative wind direction […]
153) *RWS* relative wind speed […]

Background: Fields added to IMMT-2 for VOSClim. [Note: Fields 147-153 were not included in the LMR regular section, thus future work should seek to recover this information for data that were translated into IMMA from LMR.]

158) *BMP* background (bckd.) *SLP*
159) *BSWU* bckd. wind U-component
160) *SWU* derived wind U-component
161) *BSWV* bckd. wind V-component
162) *SWV* derived wind V-component
163) *BSAT* bckd. air temperature
164) *BSRH* bckd. relative humidity
165) *SRH* (derived) relative humidity

166) *SIX* derived stn./wea. indic. (unused) [Additional information: possibly this field could be deleted from this attm, since is remains unused]
167) *BSST* bckd. *SST*
168) *MST* model surface type
169) *MSH* model height of surface
170) *BY* bckd. year
171) *BM* bckd. month
172) *BD* bckd. day
173) *BH* bckd. hour
174) *BFL* bckd. forecast length (do not use; erroneous in R2.5 data) [Additional information: This 'do not use; erroneous in R2.5 data' label will need to be removed prior to R2.6.0 as all data will be reprocessed and erroneous '99' values removed.]

Model quality control feedback information.

Background: Upon receipt of each GTS report from a VOSClim ship, the VOSClim Real Time Monitoring Centre (RTMC; at the UK Met Office) appends co-located parameters (and related information) from the Met Office forecast model for six variables—*SLP*, wind U- and V-component, air temperature, relative humidity, and *SST*—to a selection (translated into BUFR) of the originally reported GTS data. These augmented ship reports are made available in BUFR format to the VOSClim Data Assembly Center (DAC; at NOAA/NCDC), which converts them into IMMA format, including this attachment. Presently *SIX* is unused (should always be missing) because it is not among the fields in the input UK BUFR format. Beginning in June 2011, the RTMC extended the model feedback information provided to the DAC to the full VOS fleet, plus moored and drifting buoys. The DAC continues to parse the VOSClim fleet from the larger VOS set and makes these observations available (http://www.ncdc.noaa.gov/oa/climate/vosclim/vosclimdata.html). Similarly, a backlog of the VOS and buoy observations extends back to 2000 and plans are in place to receive and process those files. [Note: In R2.5 data, *BFL* was recently discovered to be subject to a conversion error and should not be used. Additionally, the original BUFR field that provides *BFL* is in minutes, thus future consideration should be given to the possibility, if appropriate, of changing the representation of *BFL* to an improved form.] [Additional information: NCDC was not reading the -9999999.0 values correctly and this field was populated as '99'. This wasn't due to the transmissions.]

Annex D: Scenarios for Adding Adjustments or QC within IVAD
(Proposed solutions to each scenario are in blue)

Scenario 1
An original air temperature record (20˚C) is adjusted for a ship heating bias (+1.1˚C) and height adjusted to 10 m (-0.2˚C). The user wants to assess the separate impacts of the bias adjustment and the height adjustment. To support the user, three separate IVAD attachments are needed.

1. All factors applied: *VAD*=20.9˚C
2. Bias correction only: *VAD*=21.1˚C
3. Height correction only: *VAD*=19.8˚C

Note: This sequence can be expanded to more than two adjustments to a single value. In many cases, storing the individual adjusted values and the all-factors-applied adjustment may be necessary. This could be complicated by order-of-operation dependent corrections (see scenario 4 below).

The author of the adjustments should submit 3 IVAD corrections (one for each case above) and document them as required. The IVAD central data center will not take responsibility for data processing of corrections in the IVAD portal. Each correction will be stored in a separate IMMA *Ivad* attm.

Scenario 2
The original *Core* air temperature (AT) record was flagged by ICOADS as out of a realistic range. A provider provides either a corrected value for *AT*, e.g. fixing transposed digits, or a bias-corrected replacement. How do we address the quality flag issue?

An error fix will go into the *Core AT* field during the next ICOADS Release and the QC flag will be set appropriately (QC is rerun with each new Release). The bias-corrected *AT* is held in an *Ivad* attm; if a user requests *AT* with all available bias corrections, the core AT—regardless of the QC, is supplied together with the bias-corrected *AT*. Thus we anticipate that the IVAD user interface will always provide the *Core* value along with available bias correction(s) whenever an adjusted value from an *Ivad* attm is requested.

In contrast, current thinking is that this user interface approach would not also apply to errors, i.e. users interested in erroneous values probably would need to get the full IMMA data set and seek out any *Error* attms.

Scenario 3
The original *AT* was in error. Provider 1 sent a corrected value. Provider 2 applies scenario 1 to the corrected value. How do we handle this with the *Ivad* attm? Would the first *Ivad* attm store the corrected value with *VEI* set accordingly, and would subsequent *Ivad* attms 2-4 store the *VAD* based on the error-corrected value of AT?

The corrected value from Provider 1 will go into the *Core* and will have a corresponding *Error* attm. Provider 2 will submit as many *Ivad* attms as necessary. By its very nature, an *Ivad* attm must reference a value in the *Core* from a specific Release (tracked via Release number details available in the *Uid* attm).

A problem would arise if Provider 2 submits an *Ivad* adjustment for an erroneous value detected either simultaneously or later in the process. A suggestion was made to track groups working on adjustments/error corrections for specific parameters, so we can avoid conflicting work. A “timestamp” may be included in the *Ivad* and *Error* attms, which may help sort out such conflicts during the processing of a new Release (see scenario 5 below).

Scenario 4
What if an adjustment actually depends on the magnitude of the value, for example, a nonlinear correction to a wind speed value because of changing ship size? Then the order of application of a bias adjustment (e.g., for ship shape, flow distortion) and subsequent height adjustment is important. Will this be handled with multiple *Ivad* attms showing the possible outcomes? The role of an expert panel (not currently funded) may be crucial to define the necessary *Ivad* attms for these unique cases.

In this case we expect there would be two *Ivad* attms. However, if this is a one-provider scenario, we recommend that the provider make a single final complete adjustment. If there are two providers, and the second uses the first *Ivad* adjusted value e.g. for flow distortion, and adds their height adjustment, the result should be a single value from provider 2. Many of these details need to be included in the provider documentation.

We also discussed establishing the Terms of Reference for an informal IVAD international coordination panel. Helping make decisions in these cases may be one role of the panel.

Scenario 5
Suppose a bias adjustment was made to an *AT* value in R2.5. However for R2.6.0, a duplicate report was received in delayed-mode. How do we deal with the duplicate and the possibility that the new and the old release *ATs* may be either identical or different? If identical, do we carry the R2.5 *VAD* temperature to R2.6.0? If different, do we scrap the *VAD* from R2.5 and notify the original provider? This is just one of many scenarios that will arise between Releases, so we need to discuss how to assign/carry information regarding applicability of each IVAD attachment for future releases.

If the original R2.5 report and the R2.6.0 duplicate contain exactly the same *AT*, it seems that the IVAD adjustment could be copied to R2.6.0 (although the mechanics of this may be challenging). We would need traceability of the *Ivad* attm back to R2.5 (tracked via Release number details available in the *Uid* attm). In addition, a "timestamp" may be included in each *Error* and *Ivad* attm.

If the R2.5 and R2.6.0 duplicate records do not have the same *AT*, it seems we must delete (i.e. no longer offer publicly, unless in the "intermediate" file product) the entire R2.5 record. The set of deleted records will then need to be analyzed. Perhaps we should state this policy in an IVAD policy document. Again, making a decision regarding the IVAD attachment from the earlier Release may be a role for the informal international coordination panel.

Annex E: Discussion of *Ivad* attm Configuration Details (w/ UK NOC)

Liz Kent and Dave Berry's comments (16 May 2012 e-mail) on an earlier version of the *Ivad* attm (in black), and responses (in blue; ref. also Shawn Smith's 13 June 2012 e-mail, providing an earlier version of responses):

1) Interplatform uncertainty. Might want to consider a more general name for this. You could have uncertainty that is correlated across e.g. platforms, measurement methods or countries. Or even for just daytime observations, or when it is raining. The flag would then have to explain how the partly correlated uncertainty linked across records.

An alternate approach would be to fix two uncertainty values to well known standards and allow the third uncertainty to be defined by the *VAD* provider, e.g. include random (*VAUR*) and bias (*VAUB*) uncertainty fields for all *VAD*. The question would be whether or not a single method for random uncertainty and bias can be established, or if an indicator field for the *VAUR* and *VAUB* would still be needed. The third "addional" uncertainty field (*VAUA*) would allow the *VAD* developer flexibility to provide any other uncertainty they feel important. A limited controlled vocabulary of uncertainty types/methods could be included in an indicator field. This field could be used for correlated uncertainties (platforms, measurement fields, etc.) or other uncertainty values that we have not considered in previous discussions.

The second alternative would be to make all three of the uncertainty fields definable, but we agreed that this could be very confusing to the users. Having at least 2 uncertainty fields that are widely used and understood (random and bias) should benefit the user community.

In conjunction with the response to question 2, we went with three *VAU* fields (as stated above for the first alternative: *VAUR*, *VAUB*, and *VAUA*), and three corresponding indicator fields (*IVAUR*, *IVAUB*, and *IVAUA*). The specific configurations for the indicator fields will need to be fleshed out during prototyping, but ideally it would be helpful if possible to agree on a limited controlled vocabulary of uncertainty types/methods to include in the indicator fields (extending e.g. to correlated uncertainties—with respect to platforms, measurement fields etc.—or to other uncertainty types that we have not considered in previous discussions).

2) I think you need an indicator for each of *VAUR/B/I* and that it has to allow for more than 0-9. For example you could imagine having different random uncertainties for each different measurement method (bucket, eri, hull, etc) and then also want to indicate how the measurement method was determined (GTS code, delayed mode, Pub. 47, estimated from country etc.) which could easily lead to many more than 10 combinations, which were differently defined for each of *VAUR/B/I*. The codes would need to be individual to each attachment (i.e. not even every SST adjustment attachment might have the same codes—although consistency would be encouraged where possible).

Agreed: to have an indicator field for each uncertainty field with more than 0-9 options. As now shown in Table C8, this is suggested to remain a 1-character field stored a base36 number.

Clearly of course not all information about the individual uncertainty values is going to be able to be captured in an indicator field. Much of the details on how the uncertainties were created and how best to use them must be captured in the supporting documentation, referenced by the *ARC*.

3) *ARC*—allow for 2 characters rather than 2 digits?

Agreed: We modified the confirmation of *ARC* (author reference code–*Ivad*), as well as that of the similar fields *ARCE* (author reference code–*Error*) and *ARCT* (author reference code–*Track*), to two (strictly alphanumeric) characters (thus disallowing symbols). As noted in the relevant tables, this also permits interpretation as base36.

This will cover the possibilities of (i) using characters rather than digits to convey meaning to *ARC* (e.g. BK might mean Berry and Kent, 2011), but also using a larger range of 2-character base36 numbers. Guidance decisions will be needed later, but tentatively we suggest ICOADS/IVAD centrally should probably assign this field (or at least its permanent value) rather than the IVAD providers.

4) Do we need a flag to say which (if any) is the recommended attachment—i.e. the default that would be obtained via the NCAR interface? This could give more information e.g. published and current; published and superseded; unpublished etc.

If we understand correctly this idea seems hard to manage and carries a presumption of recommendation that might not ubiquitously apply to all usage situations. Through the GUI, and of course in the supporting IMMA records, we will offer all *Ivad* attms. We also feel what is recommended is important and could change over time. Originally, we had thought not to seed any check box clicks in the NCAR GUI as a default, but maybe this would be helpful—we would take advice from IVAD expert teams on this. If a flag was set in the data record, in some cases it would need to be changed, e.g. if it was superseded, this would be an additional complicating step when adding new *Ivad* records.

We will continue to consider designing the columns of information in the *ARC* master table to ensure we can capture the status of the reference document.

We note that every *Ivad* attm submitted to ICOADS will be assigned a "date stamp" (*AJDN* in Table C8) for insertion into ICOADS that may be useful to support the expert teams' decisions on recommended adjustments. Also however the procedure to "unpublish" an *VAD* correction is not yet clear, i.e. policies as to when and how *VAD* corrections will be removed from an ICOADS Release.

5) iVAU. Rather than being a combined flag for *VAUR/B/I* this should probably be a link to documentation.

Agreed: As discussed under question 2, a separate indicator has been included for each uncertainty value. Again however much of the documentation for the uncertainty calculations and application will probably need to be separate and linked to via *ARC*.

6) The *VQC* is rather different to the concept of trimming—is there any idea that they might relate in any way? Also I know it's based on the MQCS flags but it might be useful to differentiate between values that "appear to be erroneous" and those that are clearly unphysical.

These are good questions. Table C8a outlines the current approach, but other proposals and refinements, as we progress through prototyping process would be welcome. The ICOADS trimming flag configuration is currently as shown in Table E1, with possible mappings to Table C8a—this illustrates your very accurate observation that *VQC* is rather different to the concept of trimming, since the mapping in Table E1 is not always straightforward or even in some cases resolvable. Table E2 shows a possible alternative approach to Table C8a in configuring *VQC*, based approximately on the proposal made to the IODE-JCOMM ODS process, which could include a value as suggested for "clearly unphysical."

Table E1. Trimming flag values (left) and possible mappings to *VQC* (right). Some mapping decisions, such as where to draw the boundary between good, suspect, and erroneous data, are subjective.

Trimming flag values	*Proposed mapping to VQC (Table C8a)*
(missing trimming flag)	0: No QC has been performed on this element
1: within 2.8σ limits	1: QC has been performed; element appears correct
2: <2.8σ sigma lower limit […][1]	"
3: >2.8σ upper limit […][1]	"
4: <3.5σ lower limit […][1]	3: QC has been performed; element appears doubtful
5: >3.5σ upper limit […][1]	"
6: <4.5σ lower limit	4: QC has been performed; element appears erroneous
7: >4.5σ upper limit	"
(8-10 unused)	(N/A)
11: limits missing (ocean/coastal box) […][2]	(N/A; flag relates to independent QC process results)
12: limits missing (ocean/coastal box)	0: No QC has been performed on this element
13: landlocked 2-degree box	(N/A; there is no landlocked flag in Table C8a)
14: data unusable (*SF*, *AF*, and *PF*, only)	(N/A?)
15: data missing or not computable	9: The value of the element is missing
(N/A; trimming is essentially univariate, except wind *U*/*V* and derived variables)	2: QC has been performed; element appears inconsistent with other elements
(N/A; trimming does not change data values)	5: The value has been changed as a result of QC

1. More precisely, "<2.8σ sigma lower limit" means: $g - 3.5{*}s_1 \le a_1 < g - 2.8{*}s_1$, where a_1 is the individual observation under scrutiny, g is the smoothed median, and s_1 and s_5 are the smoothed lower and upper median deviation (similarly for the other trimming flag values footnoted).
2. Special value for MEDS buoy data, such that MEDS flagged the data as correct (for *SST* and *SLP* only).

Table E2. Possible reorganization and expansion of the standardized flag scheme proposed to IODE-JCOMM Ocean Data Standards (ODS) (see Annex F, Table F2). This could form an alternative approach to configuring *VQC*.

Code	*Primary level flag's short name*	*Definition*
1	Good	passed documented required QC tests
2	Questionable/suspect	failed non-critical documented metric or subjective test(s)
3	Bad	failed critical documented QC test(s) or as assigned by the data producer
4	Unphysical value	*(note: not part of ODS proposal)*
7	Landlocked[1]	*(note: not part of ODS proposal)*
8	Not evaluated, not available or unknown	used for data when no QC test performed or the information on quality is not available *(note: positioned instead between good and questionable/suspect in ODS proposal)*
9	Missing data	used as placeholder when data are missing

1. One possible argument against such a flag value, which might be raised e.g. by proposers to ODS, is that setting this flag then precludes performing other standard QC tests on the data value and then assigning a value 1-4. As one related consideration in the early historical data context, reported ship positions could be far less accurate than modern navigation allows, thus near-coastal reports might be recorded in logbooks at positions that are over land according to modern standards—but the data might be usable if repositioned.

Earlier response comments (13 June): VQC is designed to be a wide-open QC field. The vision would allow an IVAD developer to create any type of advanced QC and distill the results down to this short list of QC flags. For example, if someone wanted to create a combined flagging scheme that included input from the MQCS, NCDC, and ICOADS flagging schemes, they could establish a procedure and map their combined flagging approach to the simple flag structure in *VQC*. Any methods and flag mapping would be documented via the *ARC*.

Regarding the difference between “appears to be erroneous” and “clearly unphysical”, we could add a separate flag, but note that the majority of the users want less flags, not more. Most users would not differentiate between “appears to be erroneous” and “clearly unphysical”. Most users would only used data flagged as 0, 1, or 5 in the proposed scheme, and some only those data labeled 1. For example, in the SAMOS QC, we conduct a range check for physical limits and flag data as “erroneous” just by visual inspection. We would map both these flags into 4 because the data should not be used. Whether or not we placed the former into a separate “clearly unphysical” flag would no change the “do not use” recommendation. The only reason to include a flag that states the data “are incorrect” vs. “appears to be erroneous” would be if we believe a user would treat these flags differently.

7) Regarding implementation. The 2 methods described (extending the IMMA line or linked attachments) both have their attractions. We discussed this issue and felt that the linked method probably provides the flexibility that we believe is needed. It also has practical attractions - you don't have to download the whole of ICOADS again just to get a new adjustment. It also allows researchers to use the same methodology to exchange information for testing etc. taking advantage of the software IVAD provides or adapting it if something slightly different is needed for a particular application.

Agreed: the IMMA1 design and modifications to the rdimma1 software are progressing in that direction.

8) If the field is linked rather than attached to the record then that might be a reason to provide the adjustment in VAD rather than the value.

At this stage, we propose sticking with the existing plan to include only the final adjusted value in the *Ivad* attm. Resources at this time do not allow us to consider all the ramifications of changing to include just the adjustment value. As one consideration, the adjustment values might be positive or negative, which may not fit with the tightly controlled range of the final adjusted data values, in that the control information for adjusted field characteristics might then no longer be strictly inherited from *ICN* & *FN*.

Annex F: QC Flag Discussion: Oceanographic and Marine Meteorological Quality Control Schemes; and Proposal to Adopt a Common Value-Added QC (*VQC*) Flag

Background
This Annex briefly reviews a variety of quality control (QC) flag schemes currently available from various oceanographic and marine meteorological datasets, building on the previous work of DMPA (2008), which reviewed two QC flag schemes currently used within ICOADS processing: (i) NOAA National Climatic Data Center Quality Control (NCDC-QC) and (ii) "trimming." DMPA (2008) also reviewed three selected flag schemes external to ICOADS: (iii) the JCOMM Minimum Quality Control Standard (MQCS), (iv) Shipboard Automated Meteorological and Oceanographic System (SAMOS), and (v) Global Ocean Surface Underway Data (GOSUD).

This Annex also discusses published work comparing a wide range of existing QC flag schemes, together with a recent IODE-JCOMM Ocean Data Standards (ODS) proposal for a standardized quality flag (QF) scheme. QF schemes managed by the following projects were among those reviewed for this Annex: (a) OceanSITES (2010); (b) MQCS-6 (JCOMM 2009); (c) NOAA National Data Buoy Center (NDBC) buoy and C-MAN (NDBC 2009); (d) SeaDataNet (2009); (e) Global Temperature-Salinity Profile Programme (IOC 2010); and (f) Integrated Science Data Management (ISDM) Drifting Buoys (ref. *TBD*). The Annex concludes with discussion of our recommendation for setting the Value-Added QC (*VQC*) flag in the *Ivad* attm.

Reviewing these and other QF schemes, the following conclusions by Reiner Schlitzer (http://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxnZWJpY2h3aWtpfGd4OjdhMDljMGI5NjdlMjUwNDI&pli=1) are apparent:

- Too many QC schemes currently exist
- Ranging from simple to very detailed
- Many schemes have flags that describe data history rather than data quality
- Mapping between schemes is sometimes difficult.

Mapping Between Ocean Dataset QC Flags
Published work comparing existing QC procedural results and QF schemes includes Cummings (2010) and IODE (2010). The Ocean Data View (ODV) group has provided the most comprehensive list of oceanographic QC flag schemes and their current mappings as they are used in the Ocean Data View IV to display original data points or gridded fields based on the original data from multiple and often very different datasets (ODV 2011).[2]

Standardizing Ocean and Marine Meteorological QC Flags
Following recommendations from the First IODE Workshop on Quality Control of Chemical Oceanographic Data Collections (IODE 2010), a formal proposal (Konovalov 2011) was submitted to the IODE-JCOMM Ocean Data Standards (ODS) group outlining a standardized QF scheme for all oceanographic and marine meteorological data.[3]

---

[2] ODV (2011) compared QC flag schemes (and mappings between them as implemented in ODV software) for: ODV, GTSPP, ARGO, SEADATANET, ESEAS, WOD, WODSTATION, WOCEBOTTLE, WOCECTD, WOCESAMPLE, Qartod, BODC, PANGAEA, SMHI and OceanSITES.

[3] For current status of the proposal and its community review see: http://www.oceandatastandards.org/index.php?option=com_content&task=view&id=46&Itemid=0

The following information was extracted from the proposal, outlining a two-level quality flag scheme:

- The first or primary level is composed of five quality codes and their definitions (Table F1). The second level complements the first level by reporting the results of QC tests performed and data processing history (Table F2). For example, if a data user only wants data flagged "good," then this person will only use the primary level. On the other hand, if the user needs information justifying the primary level flags, then the secondary level provides complete information on the quality test applied and their results. In this way the data user can accept or reject any data based on level 1 or make an informed choice based on level 2.
- The first level quality flags are limited in their number and restricted to those listed in Table F1. These flags are of increasing numerical value in line with the decreasing quality of data providing an easy analysis and filtering of data in a database or joining of data from different databases. The reason for a specific quality flag for a data point is justified by the results of applied quality tests, with details proclaimed at the second level. While different tests can be applied and qualified as required, the critical and non-critical tests for data sets of different nature and origin and information on the tests and their results is completely preserved at the second level. The added level of detail enables clear justification of the nature and reason of the primary quality flags.
- The second level quality flags are variable in their quantity and quality summarizing information on the applied quality tests (e.g., excessive spike check, regional data range check, etc.) and data processing history (e.g., interpolated values, corrected value, etc.). This scale makes it possible to join the gained experience and information from established programs and projects (e.g., Argo, GTSPP, OceanSites, Qartod, SeaDataNet, IMOS, MMI, WOD, etc.) and provides a possibility for additional currently unforeseen second level quality tests and procedures.

Table F1. Primary-level quality flag codes and definitions. Any quality control tests must be well documented in the metadata that accompany the data.

Code	*Primary level flag's short name*	*Definition*
1	Good	passed documented required QC tests
2	Not evaluated, not available or unknown	used for data when no QC test performed or the information on quality is not available
3	Questionable/suspect	failed non-critical documented metric or subjective test(s)
4	Bad	failed critical documented QC test(s) or as assigned by the data producer
9	Missing data	used as placeholder when data are missing

Table F2. Examples of secondary-level codes and descriptions. All objective (i.e., reproducible, numerical metric tests) or subjective (e.g., expert review) tests should be well documented, including peer-reviewed or authoritative reference sources as part of the metadata that accompanies the data.

Code	*Quality control test and data processing history*
01	regional data range check
02	excessive gradient check
03	excessive spike check
04	excessive offset/bias when compared to a reference data set
05	excessive data uncertainty
06	unexpected X/Y ratio (e.g., chemical stoichiometry or property-property X to T, S, density, among others)
07	excessive spatial pattern check ("bullseyes")
...	…
20	below detection limit of method
21	interpolated value (not measured)
22	data offset corrected value relative to a reference data
23	expert review
Etc. ...	…

Assuming the proposal is eventually adopted, how this scheme might be mandated and implemented across various data producers is still under debate and the transition may not be easy and straightforward. However, the following are listed in Konovalov (2010) as advantages for adopting this scheme:

- Small and fixed number of unambiguous flags at the primary level;
- Primary level code values are numeric and ordered such that increasing quality flag values indicate a decreasing level of quality. This supports the identification of all data that meet a minimum quality level;
- The monotonic primary scale facilitates the inheritance of quality flags for derived or calculated variables. For example, when temperature and salinity values are used to calculate density, the density value will inherit the flag of the datum with the lowest quality;
- The scheme is universal; it can be applied to all types of data making possible to merge and exchange them;
- It enables mapping between quality flags and quality tests;
- Existing QF schemes can be mapped to the proposed scheme with no information loss;
- Data sets with different QF schemes can be merged into one data set preserving all existing quality flags and making possible to apply new tests and save their result.

Planned QC flag (*VQC*) implementation for IVAD

Within the *Ivad* attm (Table C8), we envision using field *VQC* (Table C8a) as a mechanism for storing externally provided data QC information. Specifically, the provider of QC information would be requested to map their flags to the proposed 0-9 configuration for *VQC* and describe their method in external documentation as linked via *ARC* (also original QC flags could be stored in the *Suppl* attm together with original data).

For *VQC* the proposed flag scheme (Table C8a) is patterned partly after that used by the Global Collection Centers (GCCs) for the IMMT format, however 6-7 will not coincide with IMMT since those are specific to the GCCs (6-8 may be reserved for future IVAD requirements). The MCQS scheme also essentially matches the current flag scheme being used by the GTSPP group for oceanographic observations, and has some similarities to the aforementioned ODS proposal (see Table F1), as well as to some QC flag configurations defined in the BUFR format (WMO 2011).

While the goals and general advantages as described above for the primary-level QC flag scheme seem very appropriate to pursue, we decided not to use the specifically proposed Table F1 flag scheme. One particular concern we had in Table F1 was with the positioning of value 2 (not evaluated, not available or unknown), between values 1 (good) and 3 (questionable/suspect). Values 1 and 3 (also 4, bad) all imply determination of data quality via a single QC process, and clearly should appear together. In contrast, value 2 indicates that the data were not subject to the QC process, so potentially that value belongs more properly down with 9 (missing data) (see Annex H, Table H2 for a possible alternative). We further note that the IMMA format satisfies through a different approach other general goals of the ODS proposal, including that existing QF schemes can be mapped to the proposed scheme with no information loss—but IMMA does this through the preservation of original input supplementary data (including such flags) rather than a re-mapping of information to new universally defined flags.

References

Cummings, J., G. Brassington, R. Keeley, M. Martin, and T. Carval, 2010: GODAE Ocean Data Quality Control Intercomparison Project. In *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society* (Annex), Venice, Italy, 21-25 September 2009, Hall, J., D.E. Harrison, & D. Stammer, Eds., ESA Publication WPP-306 (doi:10.5270/OceanObs09) [http://www.oceanobs09.net/proceedings/ac/FCXNL-09A02-1656379-1-AC_4C_cummings.pdf].

DMPA, 2008: Common Issues of Quality Control of Surface Marine Data (Draft, 7 March 2008; Chairs: ETMC, DMCG, SAMOS, GOSUD, et al.). Doc 5.3 rev1 for *DMCG-III, Ostend, Belgium, 26-28 March 2008* [http://www.jcomm.info/index.php?option=com_oe&task=viewDocumentRecord&docID=1838].

IOC, 2010: GTSPP Real-Time Quality Control Manual, First Revised Edition. UNESCO-IOC 2010. IOC Manuals and Guides No. 22, Revised Edition (IOC/2010/MG/22Rev.) [http://www.iode.org/index.php?option=com_oe&task=viewDocumentRecord&docID=6437].

IODE, 2010: First IODE Workshop on Quality Control of Chemical Oceanographic Data Collections. Workshop Report No. 228. *5th Session of the IODE Group of Experts on Biological and Chemical Data Management and Exchange Practices, Oostende, Belgium, 8-11 February 2010* [http://www.iode.org/index.php?option=com_oe&task=viewDocumentRecord&docID=5109].

ISDM (Reference *To Be Decided*).

JCOMM, 2009: Minimum Quality Control Standard (MQCS-IV; Version 6). Annex 2 to Recommendation 9 (JCOMM-III) [http://www.wmo.int/pages/prog/amp/mmop/documents/MQCS-VI-JCOMM-III.pdf].

Johnson, D.R., T.P. Boyer, H.E. Garcia, R.A. Locarnini, O.K. Baranova, and M.M. Zweng, 2009: *World Ocean Database 2009 Documentation*. S. Levitus, Ed. NODC Internal Report 20, NOAA Printing Office, Silver Spring, MD, 175 pp. [http://www.nodc.noaa.gov/OC5/WOD09/pr_wod09.html].

Konovalov, S., H. Garcia, R. Schlitzer, L. Devine, C. Chandler, G. Moncoiffé, T. Suzuki, A. Kozyr, 2011: Proposal to adopt a quality flag scheme standard for oceanographic and marine meteorological data exchange, Version 1.1. Submitted 11 February 2011, currently under Community Review [available from http://www.oceandatastandards.org/index.php?option=com_content&task=view&id=46&Itemid=0].

NDBC, 2009: Handbook of Automated Data Quality Control Checks and Procedures, NDBC Technical Document 09-02 [http://www.ndbc.noaa.gov/NDBCHandbookofAutomatedDataQualityControl2009.pdf].

OceanSITES, 2010: OceanSITES User's Manual – NetCDF Conventions and Reference Tables, Version 1.2 [http://www.whoi.edu/virtual/oceansites/docs/oceansites_user_manual_version1.2.pdf].

ODV, 2011: Ocean Data View – Oceanographic quality flag schemes and mappings between them, Version 1.2 [http://odv.awi.de/fileadmin/user_upload/odv/misc/ODV4_QualityFlagSets.pdf].

SeaDataNet, 2009 (?): BODC Vocab Library L201 – SeaDataNet measured qualifier flags.

WMO, 2011: *Manual on Codes*, International Codes, Vol. I.2. WMO-No. 306 [http://www.wmo.int/pages/prog/www/WMOCodes/VolumeI2.html - VolumeI2].

Annex G: Edited Cloud Report (ECR) Information **(from Carole Hahn; other background info supplied by Carole is available but not included here; some cosmetic edits made below to squeeze onto 2 pp.)**

`cjh draft 080717+ {Patterned after ICOADS IMMA0 documentation.}`

Table *Cn*. **ECR attm** ("Edited Cloud Report").
Cloud variables N, NH, H, CL, CM, and CH are similar to those given in Regular Section but may be "edited" as described in this documentation.
{Q: add 'e' to Abbr to distinguish: Ne, CLe, etc?}

---------.---------.---------.---------.---------.---------.---------.---------.---------

Doc.	Len	Abbr.	Element description*			
D	2	ATTI	attm ID			Note: set ATTI=?
D	2	ATTL	attm length			Note: set ATTL=32

EECR Basic Cloud Elements (15 characters):

Doc.	Len	Abbr.	Element description	Min.	Max.	Units (Code)
?	2	ww	present weather	0	99	(ww; missing=-1)
?	1	N	total cloud amount	0	8	(N; N=9 edited)
?	2	NH	lower cloud amount	0	8	(Nh; NH=9 edited, miss=-1)
?	2	H	lower cloud base height	0	9	(h; missing=-1)
?	2	CL	low cloud type	0	11	(CL edited; missing=-1)
?	2	CM	middle cloud type	0	12	(CM edited; missing=-1)
?	2	CH	high cloud type	0	9	(CH edited; missing=-1)
?	2	CC	Change Code	0	9	(CC Table)

EECR Derived Cloud Elements (8 characters):

Doc.	Len	Abbr.	Element description	Min.	Max.	Units (Code)
?	3	AM	middle cloud amount	0	800	0.01 oktas; missing=900
?	3	AH	high cloud amount	0	800	0.01 oktas; missing=900
?	1	UM	NOL middle amount	0	8	oktas; missing=9
?	1	UH	NOL high amount	0	8	oktas; missing=9

EECR Sky Brightness Elements (9 characters):

Doc.	Len	Abbr.	Element description	Min.	Max.	Units (Code)
?	1	SBI	sky-brightness indicator	0	1	
?	4	SA	solar altitude	-900	900	0.1 degrees
?	4	RI	relative lunar illuminance	-110	117	hundredths

*Brief description of ECR variables (see H99 or H95 for details):

ECR: "Edited Cloud Report"; short for EECR *described in H99*.

ww,N,NH,H,CL,CM,CH: These **weather and cloud variables** are coded as specified by WMO except that items CL and CM have been "***extended***" *as indicated in **Table e2***. Also, cases of N=9 with fog or precipitation have been converted to N=8. Any such conversion is recorded in the "change code" (CC). *An ECR attachment is provided only if N is given in the original report*.

CC: The **change code** indicates whether the original report was changed (edited) during processing. Code values are defined in ***Table e3*** *(and **Section 3.3 of H99**)*.

AM, AH: These variables give the **"actual" cloud amounts** of middle and high clouds, derived from N and NH with use of the random overlap equation if necessary ***(Section 3.5 of H99).***

UM, UH: These variables, derived from N and NH, give the **"non-overlapped"** (NOL) **amounts** of middle and high clouds; i.e. the amounts visible from below ***(Section 3.5 of H99)***.

SBI: The sky-brightness indicator has a value of "1" (light) if the illuminance criterion described in *H95* was satisfied at the time and place of the report, suggesting that there was adequate light for visual observation of cloud cover and cloud types (if not, then SBI=0; dark). This variable can be used in lieu of SA and RI if one accepts the criterion recommended in *H95*.

SA, RI: These variables give the **solar and lunar parameters** needed to determine the illuminance provided by the sun or moon for the date, time and location of the report ***(Section 3.6 of H99).*** SA is the altitude of the sun above the horizon. RI is the relative lunar illuminance, defined in H95, which depends on the lunar altitude and phase, and the earth-moon distance. The illuminance criterion of Hahn et al. (H95) is satisfied (SBI=1) when SA≥-9° or RI>0.11. A negative value of RI means the moon was below the horizon.

ECR Table e2. **Cloud** and Weather Type **Definitions Used** in ECRs *{modified from Table 2 of H99}*

Level	Shorthand notation	Meaning	Synoptic codes	Extended ECR codes#
	TC	total cloud cover	N = 0-9	0-8
	Cr	completely clear sky	N = 0	
	Ppt	precipitation	ww= 50-75,77,79,80-99	
	D	drizzle	50-59	
	R	rain	60-69	
	S	snow	70-75,77,79	
	Ts	thunderstorm or shower	80-99	
Low			CL=	
	Fo	sky obscured by fog	/ with N=9 and ww=10-12,40-49	11
	St	stratus	6,7	
	Sc	stratocumulus	4,5,8	
	Cu	cumulus	1,2	
	Cb	cumulonimbus	3,9, or N=9 with ww=Ts	10
Mid			CM=	
	Ns	nimbostratus	2,7, or N=9 with ww=DRS	12,11,10
			/ with ww=DRS and CL=0,7	10
			/ with ww= RS and CL=4-8	10
	As	altostratus	1; 2 if not DRS	
	Ac	altocumulus	3,4,5,6,8,9; 7 if not DRS	
High			CH=	
	Hi	cirriform clouds	1-9	

Used in the EECRA dataset (H99). Extended codes are shown where they differ from synoptic codes.
In the extended code the value "-1", rather than "/", is used to signify missing information.

ECR Table e3. **Change Codes** for Edited Cloud Reports *{from Table 3 of H99}*

			Occurrence (%) *			
			Land		Ocean	
CC#	Case (brief description)	Changes made	all obs	light obs	all obs	light obs
0		none	87.4	87.4	87.2	86.9
1	N=9 with precipitation or fog	N=8; CL=10,11 or CM=10	1.6	1.6	2.6	2.6
2	Nh=0 with CM>0 and CL=0	Nh=N	0.8	0.8	0.5	0.5
3	Nh=N with CH>0 and CL=CM=0	Nh=0	0.1	0.1	0.2	0.2
4	Nh<N where it should be Nh=N	Nh=-1	0.3	0.4	0.6	0.6
5	CL =/ with CM or CH not /	CM,CH =-1	0.1	0.1	0.5	0.5
6	CM or CH miscoded as 0	CM or CH =-1	3.2	3.5	3.7	4.1
7	CM=7,2 for Ns	CM=11,12	3.7	3.5	1.1	1.2
8	CM=/ for Ns	CM=10	2.4	2.2	1.8	1.9
9	CM or CH miscoded as /	CM or CH =0	0.3	0.3	1.8	1.5

Also order in which changes are made, but CC=9 is recorded only if no previous change occurred (this conflict can occur only with CC 7 or 8).
* Data years 1982-1991.

Referenced EECR Documentation:
{H99} Hahn, C.J., and S.G. Warren, 1999: *Extended Edited Synoptic Cloud Reports from Ships and Land Stations Over the Globe, 1952-1996*. NDP-026C, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN. (Also available from Data Support Section, National Center for Atmospheric Research, Boulder, CO.)
{H95} Hahn, C.J., S.G. Warren and J. London, 1995: The effect of moonlight on observation of cloud cover at night, and application to cloud climatology. *J. Climate*, **8**, 1429-1446.